InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records

Luciana Duranti and Randy Preston

Editors

Published by
Associazione Nazionale Archivistica Italiana
Padova, Italy
2008

Status: Final (public)

Version: Electronic

Publication Date: 2008

URL: http://www.interpares.org/display_file.cfm?doc=ip2_book_complete.pdf

How to Cite: Luciana Duranti and Randy Preston, eds., *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records* (Padova, Italy: Associazione Nazionale Archivistica Italiana, 2008), electronic version. <http://www.interpares.org/display_file.cfm?doc=ip2_book_complete.pdf>

Contents

Acknowledgements

As the Director of the InterPARES 2 Project, I wish to thank first and foremost the Social Sciences and Humanities Research Council of Canada (SSHRC), which has funded the general direction of the Project and its Canadian researchers with a second MCRI grant; the University of British Columbia Vice President Research, Hampton Fund, and Dean of Arts; the United States National Historical Publications and Records Commission (NHPRC) and National Science Foundation (NSF); the Associazione Nazionale Archivistica Italiana (ANAI), and all the institutions and organizations that have provided support to the individual researchers and are identified in Appendix 1 in connection with the researcher's name. Special thanks go to the School of Library, Archival and Information Studies for providing the office space for the headquarters of the Project.

I wish to give my most heartfelt thanks to the following InterPARES Project staff for their deep commitment to the research and the researchers, and for their creative contributions to every aspect of the management of this gigantic project: University of British Columbia Project Coordinators Tahra Fung (2002-2003), Yau Min Chong (2003-2005), Greg Kozak (2005-2006) and Randy Preston (2006-2007), Technological Coordinator Jean-Pascal Morghese (2002-2007) and Project Assistant Sandy Orr (2002-2007); and State University of New York at Albany Project Coordinator Mark Wolfe (2002-2006).

Special thanks go to the Post-doctoral Fellows Brent Lee (2002), Jean-François Blanchette (2002-2004) and Bonnie Mak (2004-2006), for their invaluable contributions to the direction and content of the Project, and for their supervision of all University of British Columbia graduate research assistants throughout the Project.

I would like to thank the members of the Advisory Board, who have individually provided input and taken part in research workshops when asked. They are: Isabelle de Lamberterie, Centre National de la Recherche Scientifique (CNRS) (France); Chris Hurley, Commonwealth Bank (Australia); Shadrack Katuu, Nelson Mandela Foundation (South Africa); S. Timothy Maloney, University of Minnesota (USA); John McDonald, Consultant (Canada); and Donald Sawyer, NASA (USA).

Finally, I wish to express my deepest gratitude to the two individuals who have shared with me all the responsibilities for the management of the Project, Terry Eastwood (University of British Columbia—School of Library, Archival and Information Studies) and John Roeder (University of British Columbia—School of Music), and without whom InterPARES 2 could not have been nearly as successful as it has been.

Luciana Duranti

Introduction

The "International Research on Permanent Authentic Records in Electronic Systems (InterPARES): Experiential, Interactive and Dynamic Records," hereinafter called InterPARES 2, was carried out between January 2002 and December 2006. Its goal was to develop a theoretical understanding of the records generated by experiential, interactive and dynamic systems, of their process of creation, and of their present and potential use in the artistic, scientific and governmental sectors, and, on the basis of that understanding, to formulate methodologies for:

- ensuring that the records created using these systems can be trusted as to their content (that is, are reliable and accurate) and as records (that is, are authentic) while used by their creator;
- selecting the records that have to be kept for legal, administrative, social or cultural reasons after they are no longer needed by their creator;
- preserving the selected records in authentic form over the long term; and
- analyzing and evaluating advanced technologies for the implementation of these methodologies in a way that respects cultural diversity and pluralism.

InterPARES 2 built upon the findings of a previous research project, "The Long-Term Preservation of Authentic Electronic Records," hereinafter called InterPARES 1, which produced: (1) requirements for authenticity for electronic textual records generated in administrative environments in databases and document management systems; (2) methods for appraisal and preservation of such records; and (3) an intellectual framework for the development of preservation policies, strategies and standards. Therefore, an additional objective of InterPARES 2 was to further develop some of the findings of InterPARES 1 to test and/or extend their applicability to the new types of record-creating environments that fall under the scope of this second Project. For example, one of the primary findings of InterPARES 1 was that the chain of preservation of digital records has to begin with records creation, and this is the reason why InterPARES 2 dedicated most of its research efforts in the first half of the Project to the analysis of the creation of the records that it aims to preserve. At the same time, InterPARES 2 sought to avoid the problems incurred in the course of InterPARES 1 that resulted from that project's pre-established epistemological perspective on the concept of record. Thus, the InterPARES 2 researchers decided not to define at the outset the concept of record, instead leaving it completely open to any possibility as presented by the research findings and, consistent with this stance, to accompany the deductive approach with an inductive one.

Due to the highly interdisciplinary nature of the Project, the objectives of InterPARES 2 were addressed since the beginning with the active participation of all stakeholders: records creators (scientists, artists, government bodies, corporations and industry), members of the information technology sector, and the archival and information science and conservation professions were involved with scholars of Archival Science, Chemistry, Computer Engineering, Computer Science, Dance, Diplomatics, Film, Geography, History, Information Studies, Law, Library Science, Linguistics, Media Studies, Music, Performance Art, Photography and Theatre in the formulation and selection of case studies and general studies, the gathering of empirical evidence and analysis. This tight research collaboration was meant to ensure that the Project's results would find ready acceptance within the targeted communities. The countries actively involved in the Project were: Canada, the United States, Australia, Singapore, China, Belgium, France, Ireland, Italy, The Netherlands and the United Kingdom. In addition, a member of the Advisory Board was from South Africa; thus, researchers from five continents provided input to the research.

InterPARES 2 was led by myself, Luciana Duranti (The University of British Columbia, School of Library, Archival and Information Studies, hereinafter SLAIS), and included the participation of the researchers, research assistants, advisory board members and staff listed in Appendix 1.

The researchers belonged to one or more research units and, in the course of the Project, often moved from one unit to another as needed. The matrix in Figure 1 depicts the various individual research units, and their relationships, that comprised the InterPARES 2 Project team. Each focus and each domain comprised three working groups that conducted research jointly as a Focus Task Force or a Domain Task Force. The cross-domains were composed of representatives of each task force and of individuals whose activity was entirely dedicated to the specific area of investigation, both of whom conducted research jointly as Cross-domain Research Teams. The purpose of this complex structure was to enable the gathering of the relevant knowledge from each specific area of endeavour (focus), share it within each record related function (domain), and from this cross-fertilization build new knowledge, applicable to all areas of endeavour and expressed in common terminology, policies, descriptive schemas and models (cross-domains). To maximize the outcomes of these dynamics, each research activity within each focus, domain or cross-domain was carried out according to the most appropriate methodology for that activity as identified by the specific team responsible for it.

	FOCUS 1 Artistic activities	FOCUS 2 Scientific activities	FOCUS 3 Governmental activities
DOMAIN 1 Records creation & maintenance	Working Group 1.1	Working Group 1.1	Working Group 1.3
DOMAIN 2 Authenticity, accuracy & reliability	Working Group 2.1	Working Group 2.2	Working Group 2.3
DOMAIN 3 Methods of appraisal & preservation	Working Group 3.1	Working Group 3.2	Working Group 3.3

Terminology
Policy
Description
Modeling

Figure 1. Matrix depicting Focus Task Forces, Working Groups and Cross-domain Task Forces

Indeed, the various InterPARES 2 methodologies were inspired by the following principles:

1. Interdisciplinarity

The Project was interdisciplinary in the measure in which its goal and objectives could only be achieved through the contribution of several disciplines. For example, one of the methods chosen to develop recommendations for creating records whose accuracy and reliability can be protected over time was to conduct an exploratory study of cases in each of the areas of activity identified. To analyze the nature, characteristics, behaviour, relationships and process of creation of the interactive, dynamic and experiential records produced in the course of artistic, scientific and electronic government activities, we needed to gather a deep understanding of those activities, their purpose, their phases and the component actions, their by-products and their structure, and their context, but also their technological environment and their use. Thus, to understand the records generated in the course of producing digital music, for example, we needed music theorists and composers, as well as computer engineers and scientists, and music historians. Afterwards, to analyze the results of the case studies, we needed the contribution of methodologies, such as text analysis, diplomatic analysis and statistical analysis, developed in the context of several other disciplines.

2. Transferability

The ultimate goal of the Project was archival in nature, being concerned with the development of trusted record-making and record keeping systems and of a preservation system capable of ensuring the authenticity of the records under examination over the long term. This meant that the work carried out throughout the Project in the various disciplinary areas had to be continuously translated in archival terms and linked to archival concepts, which are the foundation upon which the systems intended to protect the records are to be designed. However, for the outcomes of the Project to be of value to each of the disciplinary areas, the findings, recommendations and archival systems envisioned by the researchers needed to be made comprehensible to the many interested stakeholders. In other words, the research outcomes had to be translated back into the language and concepts of each discipline that needed to make use of them.

3. Open inquiry

InterPARES 1 had its epistemological roots in the humanities, specifically in diplomatics and archival science. In contrast, InterPARES 2, while planning as one part of its research to test some of the outcomes of InterPARES 1 in a range of applied settings, espoused no epistemological perspective or intellectual definitions *a priori*. Instead, researchers in each working group identified the perspective(s), research design and methods that they believed to be most appropriate to their inquiry. Thus, each research team was, to a large degree, free to carry out its research using whatever methodology and tools—surveys, case studies, modeling, prototyping, diplomatic analysis, etc.—that the team considered to be the most appropriate to be able to collect and analyze the data needed to address the Project's core research questions.

The reason for this openness is that InterPARES 2 was conceived to work as a "layered knowledge" environment, in the sense that some of the research work intended to build upon knowledge developed in the course of the UBC Project[1] and InterPARES 2, some took

[1] The research project variously known as the UBC Project, the UBC-MAS Project and the UBC-DOD Project, conducted by Luciana Duranti and Terry Eastwood between 1994 and 1997, in collaboration with the Department of Defense of the United States, was concerned with the development of a record making and recordkeeping system capable of ensuring the long term preservation of authentic electronic records. One of its spin-offs was the DoD 5015.2 standard. See Luciana Duranti and Heather MacNeil (1996), "The Protection of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project," *Archivaria* 42 (Fall): 46–67. Online reprint available at http://journals.sfu.ca/archivar/index.php/archivaria/article/view/12153/13158; and Luciana Duranti, Terry Eastwood and Heather MacNeil, *Preservation of the Integrity of Electronic Records* (Dordrecht: Kluwer Academic Publishing,

knowledge of similar issues developed in other areas of endeavour and brought it to bear on records creation and preservation, some reconciled knowledge about records and their attributes, elements, characteristics, behaviour and qualities existing in various disciplines and developed it for archival purposes, and some explored new issues and studied entities never examined before, thereby developing entirely new knowledge.

4. Multi-method design

Each case study as well as each of the other research activities was carried out using the methodology and the tools that the dedicated investigating team considered the most appropriate for it. For this reason, the methodology of the Project is not presented in this book as a chapter but is discussed in the context of each activity described in each report.

The organization of the InterPARES 2 Project findings in this book reflects the structure of the research units. An integrated Focus report opens the series of the reports because it contains all the basic research upon which the work described afterwards has relied and built. The Domains' and Cross-domains' reports follow in logical order. The individual products of each research unit are attached as appendices to the end of the book. Likewise, all references cited in the individual reports are consolidated into a single bibliography attached as an appendix to the end of the book.

The appendices following the reports include several documents that are in themselves major products of the InterPARES 2 Project, such as the *Creator Guidelines—Making and Maintaining Digital Materials: Guidelines for Individuals*, the *Preserver Guidelines—Preserving Digital Records: Guidelines for Organizations* and the *Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records*. Among the many useful appendices, however, I wish to draw your attention at this point to the second appendix, which contains a reprint of an article written by Ken Thibodeau and myself on the concept of record in the environments studied by the Project as it finally emerged from all the case studies and general studies carried out in the context of the various focuses. The conclusions of the article are linked to the evidence offered by the research described in the Focus report and included in the many resources on the DVD that accompanies the book.[2] Because the article serves, in a way, as a condensed summary of the key findings of the Project as regards the concept of record in the environments studied, some readers may actually find it helpful to read the article first.

Although the content of each report is the result of the research of several co-investigators and graduate research assistants, the reports are authored by the individual researchers who took responsibility for assembling all the findings in a coherent narrative and for telling you the story of all our tribulations and successes. It is my hope that you will find this story not only instructive but interesting, challenging and stimulating as it was for us living it together day after day for five years.

Luciana Duranti
InterPARES 2 Project Director

2002). An online reprint of the book is available at http://www.interpares.org/book/index.cfm. See also the UBC Project's Web site at http://www.interpares.org/UBCProject/index.htm.

[2] All of the InterPARES 2 documents, reports and resources cited throughout the book are available on the InterPARES Web site (http://www.interpares.org/) and most are also included on the DVD that accompanies the book.

PART ONE

CASE AND GENERAL STUDIES IN THE ARTISTIC, SCIENTIFIC AND GOVERNMENTAL SECTORS

Focus Task Force Report

by

Yvette Hackett, Library and Archives Canada
William Underwood, Georgia Tech Research Institute
Philip Eppard, University of Albany, State University of New York

Introduction

The research of the Focus Task Forces constitutes the primary source of information upon which the findings and products of the Project's domains and cross-domains are based. Proper contextualization and understanding of the findings and products of the domains and cross-domains therefore requires a sufficient level awareness of the activities of the Focus Task Forces.

Background and mandate

The organizational structure of the InterPARES 2 Project was developed to address the Project's guiding methodological principles of interdisciplinarity, transferability, open inquiry and multi-method design. To achieve a high level of multicultural and interdisciplinary collaboration, the research was structured into several intersecting areas of inquiry. The research teams responsible for each area were composed of investigators from a variety of disciplines and cultural backgrounds. The research concerning the various disciplines was divided into three focuses of inquiry, each of which was tasked with examining records created in the course of one type of activity as follows:

- the Focus 1 Task Force studied records created as part of artistic activities;
- the Focus 2 Task Force studied records created as part of scientific activities; and
- the Focus 3 Task Force studied records created as part of governmental activities.

Research initiatives

As anticipated in the original research proposal, the focus task forces directed a good deal of their energy into case studies.

> The primary role of Focus Task Forces is to gather and analyze case studies and other data of relevance to each type of activity across multiple domains of inquiry.[1]

Over the course of the Project, twenty-seven case studies were proposed and approved, addressing all three focus areas. Topics included performance art, moving images and music; archaeology, cybercartography and astronomy; and taxation, Supreme Court and land records, among many others. In all, twenty-three of the approved case studies were completed.

A small number of case studies were not completed for a variety of reasons. The number of organizations participating in InterPARES 2 changed over the course of the Project, as did that of individual researchers contributing to the work. One case study was simply re-classified from a case study to a general study when the distinction between these two types of research activities became more clearly defined.[2] One case study was proposed by an organization that left the Project early. Two others were proposed by researchers who were subsequently drafted into research initiatives in the Project's Domains or Cross-domains, eliminating the time they might have had available to conduct a case study. Finally, there were three instances where the subject of the case study withdrew from the process before data collection was finished. Of these three

[1] InterPARES 2 Project, "Milestones Report," October 2002, 2 (item 1.5). Available at http://www.interpares.org/display_file.cfm?doc=ip2_milestones(200210).pdf.

[2] Case study 04, Persistent Archives Based on Data Grids, became general study 01 (under the same title).

instances, one case study was completed, one ended with an interim report and a third could not be completed.

Despite the five-year period covered by the Project,[3] it was not possible to conduct case studies into all aspects of records creation and maintenance in all the specific disciplines of the artistic, scientific and governmental sectors. In some cases, the necessary subject expertise was not represented among the researchers in the Project; in other cases, geographic obstacles would have been too expensive to overcome; and, overall, adequate human resources were frequently lacking, given the wide scope of inquiry being attempted by InterPARES 2.

In addition to case studies, the Focus Task Forces also conducted a number of general studies. Within the context of InterPARES 2, a *general study* was defined as an investigation carried out by any one of the Project's three Focus Task Forces, within its respective scope (i.e., arts, science or government), for the purpose of achieving the Focus Task Force's objectives, but which was *not related to a specific records creator*. This was in contrast to a case study, which, although also carried out by any one of the Project's three Focus Task Forces, was an investigation that focused on the records (or some portion of the records) and records management process *of a specific creator*. Thus, whereas the primary purpose of the case studies was to gather as comprehensive an understanding as possible of the *creator-specific* activities creating the records—including their purpose, their phases and the component actions, their by-products and their structure, their context, their technological environment and their use—the primary purpose of the general studies was to gather as comprehensive an understanding as possible of the above noted records-related issues, but in relation to a wider research context involving a more generalized, yet still bounded, aggregation of records creators within each Focus.

In effect, the general studies were undertaken to help the InterPARES researchers fill particular gaps in the coverage of each Focus area by the case studies. Some general studies allowed researchers to understand the degree to which a particular case study was representative of the work practices of a particular group. Thus, for example, while case study 13 was delving deeply into the work of one composer, Keith Hamel, and the technical details of one specific composition, *Obsessed Again…*, the Focus 1 researchers understood that this single case was not necessarily representative of the full range of adoption and use of digital technologies among composers and so introduced a general study consisting of a Web-based survey designed to examine the recordkeeping practices of a broader range of composers.[4] Other general study surveys looked at recordkeeping issues related to the use of GIS technology by archaeologists,[5] the use of digital technologies by photographers,[6] the functionality of government Web sites[7] and the development of data portals and repositories in the sciences.[8] Still other general studies adopted a more traditional research/report format, such as the study of Persistent Archives based

---

[3] The Project was granted a one-year extension in 2007 expressly for the purpose of disseminating its research findings, bringing to six the total number of years during which the Project was active.

[4] See Michael Longton (2004), "InterPARES 2 Project - General Study 04 Final Report: Survey of Recordkeeping Practices of Composers." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs04_final_report.pdf.

[5] See Randy Preston (2006), "InterPARES 2 Project - General Study 09 Final Report: Digital Recordkeeping Practices of GIS Archaeologists Worldwide: Results of a Web-based Survey." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs09_final_report.pdf.

[6] See Jessica Bushey and Marta Braun (2005), "InterPARES 2 Project - General Study 07 Final Report: Survey of Recordkeeping Practices of Photographers using Digital Technology." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs07_final_report.pdf.

[7] See Mark Wolfe (2003), "InterPARES 2 Project - General Study 08 Final Report: Survey of Government Web Site Interactivity." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs08_final_report.pdf.

[8] See Tracey P. Lauriault and Barbara L. Craig (2007), "InterPARES 2 Project - General Study 10 Final Report: Study of Science Data Archives/Repositories." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs10_final_report.pdf.

on Data Grid Technology[9] and the study of the selection of digital file formats in support of digital preservation activities in a number of institutions.[10]

Of the eleven proposed and accepted general studies, nine produced final reports. The remaining two studies (general study 02, Survey and Analysis of Scientific Encoding Languages for Non-Textual Records and general study 05, An Examination of the Processes to Preserve and Manage Electronic Records: Round Three at the National Archives of Australia and WGBH) were retired prior to completion, although some preliminary reports exist. Appendix 3 contains a complete list of case and general studies and their participants, including completed, retired and irregularly classified studies.

Finally, each Focus Task Force conducted bibliographic work to identify and summarize articles from within its relevant disciplines in support of the Domain 2 Task Force's research into how the concepts of accuracy, reliability and authenticity were understood in the artistic, scientific and governmental environments.

Research Methodology

Case studies

Composition of teams

Each case study team included, at a minimum, a scholar of the activity under investigation, a technology specialist, an archival expert and a graduate research assistant. As the original Project proposal explained:

> The singular nature of the research team, comprising leading scholars in archival science, the social and physical sciences, and the creative and performing arts, as well as representatives from archival institutions and other government bodies from twenty countries and five continents, will ensure that this research, so important for the preservation of our societal memory, will remain focused on 'records' rather than on all digital objects, on the preservation of their trustworthiness both as meaningful content and as records rather than on all issues related to digital preservation, and on the protection of their cultural character rather than on imposing uniform models and applications. This crucial perspective is unique.[11]

A shortage, within the Project, of both archivists and technical specialists meant that a full team was not always available. It must also be noted that a number of the graduate research assistants who joined case study and general study teams made significant contributions to the work, functioning as fully fledged researchers.

Research questions

One of the first instruments developed for use by the Focus Task Forces was a standardized set of case study research questions. The decision to use a standardized set of questions was based on the previous experience of the InterPARES 1 researchers who, having used a

[9] See Reagan W. Moore (2004). "InterPARES 2 Project - General Study 01 Final Report: Building Preservation Environments with Data Grid Technology." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs01_final_report.pdf.

[10] See Evelyn Peters McLellan (2006), "InterPARES 2 Project - General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_english.pdf. French language version available at http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_french.pdf.

[11] Luciana Duranti (2001), "International Research on Permanent Authentic Records in Electronic Systems (InterPARES): Experiential, Interactive and Dynamic Records," SSHRC MCRI InterPARES 2 Project Proposal, 412-2001, 1.1-4–1.1-5. Available at http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf.

standardized set of questions for their case studies, found that the method greatly facilitated subsequent comparative analysis.

A standardized case study questionnaire was initially drawn up in quite general, non-discipline-specific language by a small group of researchers during the June 2002, InterPARES International Team meeting in Washington, D.C.[12] The questionnaire initially consisted of eighteen questions and was intended for use by all of the Focus Task Forces. As initially conceived, the questionnaire consisted of questions that were designed to be answered by the case study subjects themselves during interviews carried out by the InterPARES researchers. The set of questions grew from eighteen to twenty-two at the next InterPARES plenary workshop in Los Angeles.[13] The additional questions addressed the research concerns of the Policy, Description and Modeling Cross-domains.

The wider scope of inquiry of InterPARES 2, relative to InterPARES 1, was reflected in the questionnaire in two important ways. First, the questionnaire was designed to be easily adapted for use in artistic, scientific or governmental environments. Second, and even more specifically, it was designed to accommodate the known vocabulary and practices of the many specialized groups the Project would be investigating; be they composers or choreographers, biologists or geomatics experts, bureaucrats or information technology personnel. However, it was later decided that this approach might limit the ability of researchers to compare the results of the case studies, both across the focuses and even within each focus.

Consequently, the questionnaire was reformulated into a more controlled set of questions, using the more rigorous terminology of archival studies, designed to be answered by the case study researchers—with the support of a common, authoritative list of Project-related terms and concepts (i.e., the InterPARES 2 Terminology Database Glossary)—using information gathered by the researchers during the course of the case study through whatever means the researchers felt would be most effective and appropriate within the context of each case study. Thus, the goal of reformulating the questionnaire into a standardized set of questions to be answered by the case study researchers, rather than as a tool to elicit information from the case study subjects, was to, in effect, "normalize" the case study findings to facilitate subsequent comparative analysis of the case study data. From this process emerged a standardized set of twenty-three research questions that would eventually form the core of the case study reports.[14]

Selection criteria

Selection criteria for case studies were kept as flexible as possible. The Call for Case Study Proposals[15] noted the lessons learned from the InterPARES 1 case studies and suggested that the most successful case studies addressed whole recordkeeping systems in the context of documented business procedures. Currently, any "whole" recordkeeping system would be a hybrid mix of analogue and digital formats, with the digital material including both entities created in digital form and digitized copies of analogue material. In judging the case study proposals, members of the International Team attempted to balance the amount of non-digital material included with the representativeness of the digital entities and the balance of digital to analogue.

[12] InterPARES 2 Project, "Proceedings of Workshop #2," National Archives and Records Administration, College Park, Maryland, USA, 20-22 June 2002. For a summary of this workshop, see http://www.interpares.org/display_file.cfm?doc=ip2_wk02_summary.pdf. The questionnaire is provided in Appendix 4.
[13] InterPARES 2 Project, "Proceedings of Workshop #3," Crowne Plaza Beverly Hills, Los Angeles, CA, USA, 17-21 September 2002. For a summary of this workshop, see http://www.interpares.org/display_file.cfm?doc=ip2_wk03_summary.pdf.
[14] See Appendix 5.
[15] See Appendix 6.

The most important selection criterion was the ability of the proposed case studies to answer the research questions formulated by each of the Project's research Domains and Cross-domains. As a result, the digital entities being proposed generally had to reside in interactive, experiential and/or dynamic systems. Although the archival interest of the InterPARES researchers was primarily in records, there was no attempt to exclude case studies that involved other digital entities that might more often be referred to as, for example, data, publications or works of art.

Each Focus Task Force also wanted to be as comprehensive as possible in studying the various disciplines in its subject area; be it the artistic, scientific or governmental sectors. At the same time, InterPARES 1 had demonstrated the usefulness of investigating similar systems, such as student record systems in universities or patent granting institutions in various jurisdictions. These parallel studies tend to highlight multiple approaches to similar recordkeeping and preservation situations.

Finally, the evaluation of case study proposals also factored in the logistics of the case study, including geographic location, the availability of archival, technical and subject specialists to form a team and the interest of the person or organization being studied in participating.

Despite this long list of considerations, selection of case studies was primarily guided by flexibility. In fact, regarding the process for accepting or rejecting case studies, it was decided during discussions at the second InterPARES 2 workshop "that there would be no specific selection criteria," and, hence, "no grounds for rejection of proposals, only for refinement of proposals."[16]

Additional documentation

When available, case study investigators collected system documentation generated by the records creators. These ranged widely across the interests of the Domains and Cross-domains, including work procedures, policies governing creation and use, technical specifications and metadata standards.

Performances and viewings

Usually, the work processes being examined in case studies are accessible to only the small group of researchers directly involved in the work. In InterPARES 2, a number of products that were the object of case studies could be more widely presented, however.

Among those works related to the artistic activities being investigated by Focus 1, CD and DVD products, such as the 3D reconstruction of the *House of Julius Polibius* (case study 09(1), Altair4 di Roma) and the documentation of Arbo Cyber, théâtre (?) (case study 01) were presented during plenary workshops. As well, a number of InterPARES 2 researchers were able to attend a performance of *Waking Dream* (case study 15) when it was mounted in Vancouver, and a "resurrected" version of the 1992 electroacoustic composition *Obsessed Again*... (case study 13)—a work that had fallen victim to technological obsolescence—was actually performed during the February 2006 plenary workshop in Vancouver.[17]

In the sciences, the *Antarctic Treaty Searchable Database* (case study 12) can be accessed via the Internet,[18] as can the *Archives of Ontario Web Exhibits* (case study 05) studied by Focus 3.[19]

[16] InterPARES 2 Project, "Proceedings of Workshop #2," op. cit., 6.

[17] Further details about the *Obsessed Again*... resurrection experiment are provided in the section titled "A strategy for preventing technological obsolescence of an artistic work" in the Domain 2 Task Force Report. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_3_domain2_task_force.pdf.

[18] Available at http://aspire.tierit.com/. It is also noted that, although the first phase of the Cybercartographic Atlas of Antarctica (case study 06) is now accessible via the Internet (see https://gcrc.carleton.ca/confluence/display/GCRCWEB/CAA+Project+Description), the Atlas was only in a pre-Internet development phase during the course of the InterPARES 2 case study.

[19] Available at http://www.archives.gov.on.ca/english/exhibits/.

General studies

The InterPARES 2 general studies utilized several different research methodologies. For example, the arts focus and science focus task forces conducted several large-scale, Web-based, questionnaire surveys.[20] Other general study research methodologies included literature reviews, large-scale, Web-based, researcher surveys (i.e., surveys in which the researchers collected data about the subjects or organizations being investigated from readily-available information sources, often without any direct input from the subjects or organizations themselves),[21] collaborative research and interviews,[22] and tool-building and experimentation.[23]

Composition of teams

Each general study team included, at a minimum, a scholar of the general activity, discipline or community under investigation and a graduate research assistant. As noted earlier, a number of the graduate research assistants who joined the general study teams made significant contributions to the work, functioning as fully fledged researchers.

Selection criteria

Aside from the absence of a formal Call for General Study Proposals and a requirement that the studies not focus on specific records creators, the selection criteria and selection process for the general studies closely mirrored that noted above for the case studies.

Literature reviews

There was an early assumption that a review of existing literature about digital practices in the artistic, scientific and governmental sectors would also inform the work of the focus task forces. Researchers from each of the three focus task forces reviewed over 200 literature resources identified by the Project's graduate research assistants. The publications varied from product announcements, to detailed technical explanations of specific technologies, to high-level planning documents meant to precede the acquisition and use of any technology. The researchers soon discovered that, overall, very little was written about digital practices in the artistic, scientific and governmental sectors, whether from an archival perspective, or from a basic preservation perspective, or even from a practical perspective. As a result, this methodology was abandoned and researchers in each focus task force restricted their literature searches and reviews to works that discussed the concepts of authenticity, reliability and accuracy from the perspective of the researchers' individual specializations; be it in the artistic, scientific or governmental sectors. Any useful citations were forwarded to the Domain 2 Task Force, which was responsible for the analysis of these concepts.

---

[20] See Longton, "General Study 04 Final Report," op. cit.; Bushey and Braun, "General Study 07 Final Report," op. cit.; and Preston, "General Study 09 Final Report," op. cit.

[21] See Wolfe, "General Study 08 Final Report," op. cit.; Lauriault and Craig, "General Study 10 Final Report," op. cit.; and McLellan, "General Study 11 Final Report," op. cit.

[22] See Jennifer Douglas (2006), "InterPARES 2 Project - General Study 03 Final Report: Preserving Interactive Digital Music - The MUSTICA Initiative." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs03_final_report.pdf.

[23] See William Underwood and Sheila Isbell (2007), "InterPARES 2 Project - General Study 06 Final Report: A Bayesian Belief Network: Supporting the Assessment of the Degree of Belief that a Recordkeeping System Maintains Authentic Digital Records." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs06_final_report.pdf.

Analysis tools

Modeling

The final step in many of the case studies involved the creation of a "function model,"[24] using the IDEF modeling technique and IDEF0 graphical modeling language,[25] to provide a structured, graphical representation of the workflows and business practices documented in the case study final reports. The diagrams in these models illustrated the processes of creation, maintenance, use and occasionally even preservation that had been applied to the digital entities being studied.

In addition to helping the case study researchers better understand the records creation, maintenance and, where present, preservation activities of the case study subjects, this activity also resulted in case study models that could later be compared to the InterPARES 2 Chain of Preservation (COP) and the Business-driven Recordkeeping (BDR) models being developed by the Modeling Cross-domain. Such analysis could reveal deviations in practices by the case study subjects from the "ideal" models developed by the Modeling Cross-domain. These differences could suggest areas where the adoption of new procedures by the creator might ensure the authenticity of digital records and/or improve the likelihood of their long-term preservation. Alternatively, a consistent deviation in work-flows by the case study subjects in relation to either the COP or the BDR model might suggest a disconnect between archivists' perception of implementable work practices and the reality faced by the records creators.

Diplomatic analysis

Once each case study was completed and a draft report issued, graduate research assistants in the School of Library, Archival and Information Studies at the University of British Columbia performed a diplomatic analysis of each type of digital entity identified in each case study.[26] Diplomatics is a science that encompasses a set of principles and terminology that have been used to analyze records since the 17th century.[27] A significant amount of work was conducted during the InterPARES 1 Project to adapt the traditional tenets of diplomatics to the digital environment. The characteristics of a record are itemized in the InterPARES 1 Project's *Template for Analysis*[28] and discussed in the Final Report of the Authenticity Task Force.[29]

---

[24] Defined as "a structured representation of the functions, activities or processes within [a] modeled system or subject area" (United States Secretary of Commerce, *Draft Federal Information Processing Standards Publication 183*, 21 December 1993. Available at http://www.idef.com/pdf/idef0.pdf.).

[25] The IDEF (Integrated Definition for Function) modeling technique and the IDEF0 (Integrated Definition language 0) was used by both the Preservation Task Force of InterPARES 1 and the Modeling Cross-domain of InterPARES 2. A more detailed description of the modeling methodology and language within the context of InterPARES 2 research is provided in the Modeling Cross-domain Task Force Report. For a brief synopsis of the IDEF0 modeling process and language, see Randy Preston (2007), InterPARES 2 Project - Integrated Definition Function Modeling (IDEFØ): A Primer." Available at http://www.interpares.org/display_file.cfm?doc=idef0_primer.pdf.

[26] The "Diplomatic Analysis" report for each case study is available on the InterPARES Web site (http://www.interpares.org/) and is also included on the DVD that accompanies this book.

[27] For a general introduction to diplomatics, see Luciana Duranti, *Diplomatics: New Uses for an Old Science* (Lanham, Maryland and London: The Scarecrow Press in association with the Society of American Archivists and the Association of Canadian Archivists, 1998). For a discussion of diplomatics in the context of the InterPARES 1 and 2 research, see Luciana Duranti and Kenneth Thibodeau (2006), "The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES," *Archival Science* 6(1): 13–68, especially 15–21 and 52–55 (Note: a reprint of the Duranti/Thibodeau article is provided in Appendix 2. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_02.pdf.).

[28] Authenticity Task Force (2000), "Appendix 1: Template for Analysis," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 192–203. Online reprint available at http://www.interpares.org/book/interpares_book_j_app01.pdf.

For the purposes of InterPARES 2, the diplomatic analysis was largely restricted to the testing of each type of digital object against the five necessary characteristics of a record to determine if each object could be considered a record, or whether an object was more appropriately identified as, for example, data, documents or publications. Non-records generally require a simpler preservation model because they exist autonomously from other documents and their purpose is, typically, limited to dissemination of information. Briefly, to be considered a record, a digital object must:

- possess a fixed form and stable content affixed to a stable medium;
- participate in an action;
- possess an archival bond, which is the relationship that links each record to the previous and subsequent record of the same action;
- involve at least three persons: the author, addressee and writer; in the digital environment, there are two more necessary persons: the creator and the originator; and
- possess an identifiable context (i.e., the framework in which the action in which the record participates takes place), including juridical-administrative, provenancial, procedural, documentary and technological contexts.[30]

The results of the diplomatic analyses carried out on the InterPARES 2 case studies are analyzed in the report of the Domain 1 Task Force, which focuses on the creation and maintenance of records.

Terminology

To help ensure terminological consistency from one case study report to the next and to help keep the reports clear and accessible in the multi-disciplinary environment of the Project, each case study team regularly consulted the InterPARES 2 Project Glossary and Dictionary; two of three terminological instruments comprising the InterPARES 2 Terminology Database developed by the Terminology Cross-domain. As appropriate, subject-specific uses and narrow definitions of terms used in the case study reports were submitted to the Terminology Cross-domain Task Force for analysis and possible inclusion in the Terminology Database.[31]

Validation of case studies

The findings of each case study were first tabled for discussion and validation by the members of the relevant focus task force. The primary goal of these discussions was two-fold: (1) to ensure the completeness and accuracy of the findings and (2) to identify, clarify and compare similarities and differences among discipline-specific practices within the broader artistic, scientific and governmental communities of practice. Cross-community comparisons were then carried out by presenting the completed and validated case studies to the plenary sessions held during the Project's bi-annual research workshops, which all Project researchers were expected to attend regardless of whether their research was concentrated in a focus, a domain or a cross-domain task force.

---

[29] Heather MacNeil et al., "Part One - Establishing and Maintaining Trust in Electronic Records: Authenticity Task Force Report," in Duranti, *Long-term Preservation*, op. cit., 19–65. Online reprint available at http://www.interpares.org/book/interpares_book_d_part1.pdf. Diplomatics is discussed in several places in this report, especially at pp. 22–25, 33–37 and 52–56.

[30] For further elaboration on these characteristics, see the *Diplomatic Analysis Template* in Appendix 7. See also MacNeil et al., "Authenticity Task Force Report," op. cit.

[31] For further information about the Terminology Cross-domain and the InterPARES 2 Terminology Database, see the Terminology Cross-domain Task Force Report. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_8_terminology_task_force.pdf.

An additional benefit of this approach was to provide members of the Domain and Cross-domain Task Forces with early indications of what each of the case study teams was finding. Another key benefit of these large-scale discussions was that they often led to a clearer and more refined understanding of the intent of the core twenty-three research questions. Finally, this validation and discussion process often highlighted potential issues, disciplines or communities of practice requiring additional research and/or particular attention from researchers involved in ongoing case studies as well as those researchers hoping to use the case study findings to support the work of the domains and cross-domains.

Case study reporting framework

Given the great variation among the case studies undertaken in the artistic, scientific and governmental sectors, it was decided that a standardized reporting framework was necessary to "normalize" the findings and thus facilitate cross-community comparisons and use of the findings by the Project's Domain and Cross-domain Task Forces. Among other things, the use of a standardized reporting framework would assist the InterPARES 2 researchers in navigating the findings and would help highlight the similarities and differences being uncovered in the practices of the various types of records creation environments. It was also felt that a standardized reporting framework would assist anyone who might choose to consult the case studies during future research.

A standardized reporting framework was initially developed by the Focus 1 Task Force and subsequently adopted by the other two Focus Task Forces.

The standardized reporting framework included the following eight sections:

A. Overview
B. Statement of methodology
C. Description of context
D. Narrative answers to the core research questions
E. Narrative answers to applicable domain and cross-domain research questions
F. Bibliography of relevant material, including articles about the methods and works of the subject(s)
G. Glossary of terms
H. Preliminary model

It was anticipated that all case studies would, at a minimum, address sections A through D, while contributions to sections E through H would vary, depending on the nature of the case study and its findings. Each of these sections is described in greater detail in Appendix 8.

Prototyping solutions

Two early case studies offered the possibility to pursue the lessons learned in the case study beyond the final report. One such case study, VanMap (case study 24), was conducted by the government focus. The situation of the records creator in this study, which involved the use of a system (specifically, a Graphic Information System or GIS) that was configured and used in such a way that made it difficult, if not impossible, for the system to create records, was already familiar to the InterPARES 2 researchers who had also participated in InterPARES 1. Technical experts at the San Diego Supercomputer Center undertook an analysis of the system and proposed a system re-design that would ensure that the GIS could produce records and maintain them over the long term. The solution involves the preservation of historical data from the VanMap GIS in an application environment that allows "point-in-time reconstruction" of VanMap views by re-assembling the data components that were in use at the time specified in a query.[32]

The second case study where a solution could be prototyped was *Obsessed Again...* (case study 13). Amid problems of hardware and software obsolescence, the essential issue to address was that the score for the electroacoustic composition that was the focus of this study did not fully express the composer's intentions with regard to the performance of the work. An experiment in generating a better description of the work in performance, as an element of metadata, and then using those metadata to re-program the work onto a current technical platform, was attempted. The results of this experiment are described in an appendix to the case study's final report.[33] The section of the Domain 2 Task Force Report titled "A strategy for preventing technological obsolescence of an artistic work" also summarizes the case study, the metadata and the subsequent attempts to re-perform the piece in a manner acceptable to the composer.

Dissemination activities

The case studies, and some of the general studies, provided material of great interest to a number of specialized groups and associations whose users were involved in the same activities or who used the same digital technologies. These included musicians creating electroacoustic music, social scientists and government employees using Geographic Information Systems, and both producers and archivists working in the moving image and sound industries.

A sample of the target audiences to whom members of the Focus Task Forces gave presentations includes:

- American Historical Association (AHA)
- American Institute for Conservation of Historic & Artistic Works
- American Society for Information Science and Technology (ASIS&T)
- Archives Association of British Columbia
- Archives Association of Ontario
- Associacao Portuguesa de Bibliotecarios, Arquivistas, e Documentalistas (BAD)
- Association of Canadian Archivists
- Association of History and Computing UK
- Association of Italian Archivists (ANAI)
- Association of Library and Information Science Educators (ALISE)
- Association of Moving Image Archivists
- Association of Pacific Rim Universities (APRU)
- Association of Records Managers and Administrators (ARMA)
- Association for Computers and the Humanities (ACH)
- Association for Literary and Linguistic Computing (ALLC)
- Canadian Cartographic Association
- Centre de recherche interdisciplinaire en technologies émergentes (Canada)
- Centro de Estudos Judiciários do Conselho da Justiça Federal (Brazil)
- Committee on Data for Science and Technology (CODATA)

[32] For further elaboration concerning the details of this proposed preservation strategy, see Glenn Dingwall et al. (2008), "From Data to Records: Preserving the Geographic Information System of the City of Vancouver," *Archivaria* (in press).
[33] See J. Scott Amort (2004), "InterPARES 2 Project - Case Study 13 Final Report: *Obsessed Again....*" Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs13_final_report.pdf.

- Computer Applications and Quantitative Methods in Archaeology
- Daniel Langlois Foundation for Art, Science and Technology
- Digital Arts Network (DAN)
- European Commission
- Federal Institute of Access to Public Information (IFAI) (México)
- Health Canada
- Health Information Management Association of Australia
- Information Processing Society of Japan
- International Academy of Law and Mental Health
- International Association for Music, Libraries, Archives and Documentation Centres (IAML)
- International Association of Sound and Audiovisual Archives (IASA)
- International Congress on Archives (ICA)
- International Federation of Library Associations and Institutions (IFLA)
- International Musicological Society (IMS)
- Joint Committee on Antarctic Data Management (JCADM)
- Knowledge Management Africa
- le Laboratoire des nouvelles technologies de l'image, du son et de la scène at Université Laval
- le projet de Documentation et conservation du patrimoine des arts médiatiques
- Manitoba Library Association
- Manitoba Archival Association
- National Archives of Brazil
- National Archives of Mexico
- National Archives of the Netherlands
- National Archives of Singapore
- National Archives of Vietnam
- National Association of Government Archives and Records Administrators
- Preservation and Access for Electronic College and University Resources (ECURE)
- Records Management Association of Australasia
- Renaissance Society of America
- Scientific Committee on Antarctic Research
- Society for the History of Authorship, Reading and Publishing (SHARP)
- Society of American Archivists
- Special Libraries Association
- State Archives of Bologna (Italy)
- State Archives of Florence (Italy)
- State Archives of Genova (Italy)
- State Archives of Milan (Italy)
- State Archives of Torino (Italy)
- The Royal Society of Edinburgh
- UNESCO Memory of the World Program[34]

[34] A complete list of dissemination activities is available at http://www.interpares.org/ip2/ip2_dissemination.cfm?proj=ip2.

Focus 1 – the Arts

Research team

The group of co-investigators conducting research on the records of artistic activities comprised experts in music, dance, photography, e-literature, theatre, film, multimedia, visual art and the design of multimedia software and hardware. Many have taken on multiple roles in their chosen discipline, including working as practicing creative artists, theorists, historians and archivists. This rich mix of knowledge and experience offered a balance between theory and practice, as well as illustrating the potential for tension between the interests of the artists creating the records and the archivists entrusted with their preservation.

The following is a list of researchers and research assistants who contributed to the work of the Focus 1 Task Force throughout the Project:[35]

Chair:

Yvette Hackett	Jan 2001 - Dec 2006

Researchers:

Howard Besser	New York University, USA—Working Group 3.1
Marta Braun	Ryerson University, Canada—Working Group 2.1
Ann Butler	New York University, USA—Working Group 3.1
Martine Cardin	Université Laval, Canada—Working Group 1.1
Henry Daniel	Simon Fraser University, Canada—Working Group 1.1
Sidney Fels	The University of British Columbia, Canada
Yvette Hackett	Library and Archives Canada—Working Group 3.1
Keith Hamel	The University of British Columbia, Canada
Sally Hubbard	Getty Institute, USA—Working Group 3.1
Mary Ide	WGBH, USA—Working Group 3.1
Ian Lancashire	University of Toronto, Canada—Working Group 2.1
Brent Lee	Windsor University, Canada—Working Group 2.1
Michael Longton	University of Victoria, Canada—Working Group 1.1
Randal Luckow	Turner Broadcasting, USA—Working Group 3.1
Michael J. Murphy	Ryerson University, Canada—Working Group 1.1
Isabella Orefice	Associazione Nazionale Archivistica Italiana, Italy—Working Group 3.1
Jesse Read	The University of British Columbia, Canada
Andrew Rodger	Library and Archives Canada—Working Group 1.1
John Roeder	The University of British Columbia, Canada—Working Group 2.1
James Turner	Université de Montréal, Canada—Working Group 3.1

Research Assistants:

Melissa Adams	The University of British Columbia, Canada
Scott Amort	The University of British Columbia, Canada
Jessica Bushey	The University of British Columbia, Canada

[35] Researcher membership in Focus 1 changed somewhat over the five years of the Project. Among those who were interested in Focus 1 issues but were unable to participate for the full length of the Project are: Paolo Buonora, Archivo di Stato, Italy; Ben Howell-Davis, Davis International Associates, USA; Susan Kennard, Banff New Media Institute, Canada; Geoffrey Rockwell, McMaster University, Canada.

Natalie Catto	The University of British Columbia, Canada
Seth Dalby	The University of British Columbia, Canada
Heather Dean	The University of British Columbia, Canada
Glenn Dingwall	The University of British Columbia, Canada
Jennifer Douglas	The University of British Columbia, Canada
Coby Falconer	The University of British Columbia, Canada
Shanna Fraser	The University of British Columbia, Canada
Stephen Gage	The University of British Columbia, Canada
Nadine Hafner	The University of British Columbia, Canada
Peggy Heger	The University of British Columbia, Canada
Janine Johnstone	The University of British Columbia, Canada
Greg Kozak	The University of British Columbia, Canada
Tracey Krause	The University of British Columbia, Canada
David Litke	The University of British Columbia, Canada
Yvonne Loiselle	The University of British Columbia, Canada
Luke Meagher	The University of British Columbia, Canada
Cara Payne	The University of British Columbia, Canada
Philippe Perron	Université Laval, Canada
Carolyn Petrie	The University of British Columbia, Canada
Claudette Rocan	The University of British Columbia, Canada
Rebecca Russell	The University of British Columbia, Canada
Vincent Schillaci-Ventura	The University of British Columbia, Canada
Wendy Sokolon	The University of British Columbia, Canada
Adele Torrance	The University of British Columbia, Canada
Jill Teasley	The University of British Columbia, Canada
Melanie Wallace	The University of British Columbia, Canada
Keum Hee Yu	The University of British Columbia, Canada
Sherry Xie	The University of British Columbia, Canada

Selection of Focus 1 case studies

The scope of inquiry for Focus 1 was limited to an examination of records generated in the course of artistic activities. Focus 1 profited considerably from the fact that issues related to music using digital technologies had already been introduced during the first phase of the InterPARES Project.[36] This advantage allowed Focus 1 to move through the case study process ahead of the scientific and government focus groups. As a result, Focus 1 influenced methodologies and reporting structures for the whole Project.

Most of the ten case studies completed by Focus 1 involved artistic creators; both individuals and organizations. Of particular note was the investigation into digital moving image production (case study 09, Digital Moving Images - Inputs, Processes and Outputs), which eventually studied four different animation and live-action production environments, including commercial

[36] See, for example, Brent Lee and John Roeder, "The Challenge of Digital Musical Archives: Research within the InterPARES Project," in *Digital Resources for the Humanities 2001-2002*, J. Anderson, A. Dunning and M. Fraser, eds. (London: Office for Humanities Communication [King's College], 2003), 171–180; and Brent Lee (2001), "The Growing Complexity of Music Preservation," in *Preserving Authentic Electronic Records: Preliminary Research Findings, an International Symposium held February 17, 2001 at the University of British Columbia, Vancouver, Canada,* Luigi Sarno, ed., 32–40. Available at http://www.interpares.org/display_file.cfm?doc=ip1_symposium_2001.pdf.

and non-profit enterprises in Canada, Italy and the United States. Another fruitful study involved the pairing of a case study and a general study to investigate the use of digital technology by composers (case study 13, *Obsessed Again...* and general study 04, Survey of Recordkeeping Practices of Composers).

Even in Focus 1, which was the focus group with the largest membership, all sectors of the arts were not represented. Furthermore, it was not always possible to find artists willing to have their recordkeeping practices studied. In addition, some of the experts in the artistic disciplines who initially joined InterPARES 2 were unable to sustain their involvement over the five years of the Project, while others decided to concentrate their efforts on the cross-domain research areas of Policy, Description, Terminology and Modeling.

Of the two case studies that could not be completed by Focus 1, the first (case study 16, Model for Description and Preservation of Documents Created Using Unstable and Variable Artistic Techniques) focussed exclusively on metadata issues and was referred to the Description Cross-domain for possible follow-up. The second incomplete case study (case study 22, Electronic Café International: Aging Records from Technology-based Artistic Activities) was stopped when the interviewees had to withdraw from the process for personal reasons. In this case, an attempt was made to complete the final report with the information available; however, its findings could not be validated because no follow-up confirmation with the creators was possible.

Summary of Focus 1 case studies

Contemporary art practices are increasingly interdisciplinary in nature. As a result, many of the case studies undertaken as part of Focus 1 cover more than one artistic discipline. The major artistic disciplines involved in the Focus 1 case studies can generally be categorized as follows:

- Performance art
- Experimental theatre
- Dance
- Music (composition and performance)
- Visual arts including installed works with digital moving image elements
- Experimental film/video art
- Computer-generated animation
- Documentary film production

The following brief descriptions of the Focus 1 case studies identify the digital entities that were examined, while also highlighting some of the more salient findings from each study. For a more detailed contextual analysis of these case studies in relation to the records creators and the nature of the activities resulting in document creation, see the “Characterization of the Case Studies” section in the Domain 1 Task Force Report. See Appendix 3 for a complete list of participants responsible for each case study in this focus.

Case study 01: Arbo Cyber, théâtre (?)[37]

The focus of this case study is a privately-owned theatre company, *Arbo Cyber, théâtre (?)* (Arbo), located in Québec City in Québec, Canada. Arbo’s artistic output involves performing arts, visual arts and media arts. The digital entity studied was Arbo’s Web site (the

[37] See Martine Cardin (2004), “InterPARES 2 Project - Case Study 01 Final Report: Arbo Cyber, théâtre (?).” Available in English at http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_report_english.pdf, and in French at http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_report_french.pdf.

Ludosynthèse),[38] which was developed to preserve the history and memory of the troupe—in part, through the digitization of Arbo's past artistic works—while also allowing and encouraging continued audience interaction with Arbo's past works. The case study researchers follow the processes of digitization and transformation of the creative materials originally created in analogue form by Arbo.

This case study offers an opportunity to observe how Arbo's original creative materials have been appraised by the creators themselves, digitized, and integrated into Arbo's "Web site of memory." On a more general level, this case study explores and documents issues related to the digitization of analogue material as a strategy for long-term preservation by showing how selected archives can be modified by a process of electronic marking, or annotation, while still remaining linked to analogue documents.

Case study 02: Performance Artist Stelarc[39]

This case study examines a performance artist, Stelarc, who combines robotics and technology with the human form, typically in collaboration with computer programmers, technicians and scientists. The digital entity studied as part of the case study was Stelarc's Web site. Created for promotion, advertising and documentation of his performances, Stelarc also uses the Web site as an online archive of his work and performance documentation. One of the key issues examined in this case study is defining where record creation begins and ends within the context of performance art. In addition, issues related to the challenge of assessing the reliability and authenticity of the artist's records, particularly in light of the "fragility" of the environments in which the works are created and performed, are examined.

Case study 03: *Horizon Zero*/ZeroHorizon Online Magazine[40]

This case study examines a media and visual arts institute situated within an institutional context in the city of Banff, Alberta in Canada. The digital entity studied as part of this case study is the bilingual, multimedia, online magazine, *HorizonZero*, and its related database, ZeroHorizon. In particular, the case study focuses on the digital entities that are created while producing an issue of *HorizonZero*—from conception to publication—including: administrative documents (i.e., those that facilitate communication and control of the production process); journalistic/artistic documents; and programming code that creates the presentation and interactive features of the online magazine's Flash Web site.

Case study 09: Digital Moving Images - Inputs, Processes and Outputs

This study involves four separate case studies, each of which examines a different moving image production environment. Given the complexity of the moving image production process, the analysis in each case study is limited to a sub-set of digital entities generated during the production process. In particular, each case study describes and explains the processes involved in creating digital moving image products, identifies the digital by-products created at each stage of the process (i.e., pre-production, production and post-production), models the structure of each digital entity and shows the relationships between the digital entities and their relationships to the overall production.

---

[38] See http://www.lit.ulaval.ca/arbocyber/index_content.htm.
[39] Henry Daniel and Cara Payne (2004), "InterPARES 2 Project - Case Study 02 Final Report: Performance Artist Stelarc." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs02_final_report.pdf.
[40] See Brent Lee (2004), "InterPARES 2 Project - Case Study 03 Final Report: *HorizonZero*/Zero Horizon Online Magazine and Media Database." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs03_final_report.pdf.

The four production environments, and the basic production processes/digital entities examined, include:

1. a small commercial production company based in Europe (Altair4 di Roma) that produces computer-based 3D animation of historical and archaeological sites;[41]
2. a state-funded film and video production and distribution outlet (The National Film Board of Canada) and the work of its animation studio;[42]
3. a commercial animation studio and the procedures governing the production of image but not sound elements;[43] and
4. a large, U.S.-based public television production company (WGBH in Boston) and its transition from using a control system for analogue film footage to a digital asset management system.[44]

Case study 10: *The Danube Exodus*[45]

This case study examines the processes, products and by-products involved in creating *The Danube Exodus: The Rippling Currents of the River*, an interactive, multimedia gallery installation by Hungarian filmmaker and artist Péter Forgács in collaboration with an art collective (the Getty Research Institute). The digital entity studied as part of the case study is the "work," which was displayed at the Getty Research Institute in August and September 2002. As noted in the overview to the case study report, the work has a complicated history and provenance, and is (or was), effectively, a manifestation of several related pieces drawing on much of the same original material. Moreover, the work actually exists (or existed) in two forms: a gallery installation and a Web site. The gallery installation is no longer active but the Web site is still live.[46] This case study offers the opportunity to examine issues of reliability and authenticity within a virtual and hybrid context.

Case study 13: *Obsessed Again…*[47]

This case study examines *Obsessed Again…,* an electroacoustic work for solo bassoon involving the use of interactive electronics commissioned in 1992 by American bassoonist, Jesse Read, and created by Canadian composer, Keith Hamel. The digital entities studied are the instructions (score and software) required to perform the work. Although the commercial hardware and software elements used in the original composition were state-of-the-art, most have since become largely obsolete, making it difficult, if not impossible, to perform the piece today. The case study team set out to resurrect the work through a complex, controlled migration process that involved identifying both digital and non-digital documents associated with the work, articulating the requirements for musical authenticity based upon the documents, building

[41] See Isabella Orefice (2004), "InterPARES 2 Project - Case Study 09(1) Final Report: Digital Moving Images - Altair4 di Roma, A Multimedia Archaeological Project: *The House of Julius Polybius*." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-1_final_report.pdf.

[42] See Andrew Rodger (2006), "InterPARES 2 Project - Case Study 09(2) Final Report: Digital Moving Images - National Film Board of Canada." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-2_final_report.pdf.

[43] See James Turner et al. (2004), "InterPARES 2 Project - Case Study 09(3) Final Report: Digital Moving Images - Commercial Film Studio." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_final_report.pdf.

[44] See Mary Ide (2005), "InterPARES 2 Project - Case Study 09(4) Final Report: Digital Moving Images - WGBH Boston." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-4_final_report.pdf.

[45] See Sally Hubbard (2006), "InterPARES 2 Project - Case Study 10 Final Report: *The Danube Exodus*." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs10_final_report.pdf.

[46] See http://www.danube-exodus.hu/en/.

[47] See Amort, "Case Study 13 Final Report," op. cit.

a performable, authentic realization of the work and developing a method for the future storage, retrieval, migration and access of the work.

To help assess the representativeness of the findings from this case study, Focus 1 researchers also conducted general study 04 (Survey of Recordkeeping Practices of Composers), which is summarized later in this chapter.

Case study 15: *Waking Dream*[48]

This case study examines *Waking Dream*, a multimedia, performance-based art piece involving dance/movement, soundtrack, live and pre-recorded video and remote controlled interactions between performers and various digital and analog technologies. The piece is the result of the collaboration between three partners at the Human Communications Technologies (HCT) Laboratory at the University of British Columbia. The digital entities studied as part of the case study are the Web site and the various digital components that participate in the performance of the work.

Similar to the situation noted by the researchers in case study 13, much of the hardware and software originally used to stage this piece are now obsolete. Unlike the situation in case study 13, however, no records of the original *Waking Dream* performance exist in any form,[49] making the task of staging an accurate and authentic reproduction of the piece all the more tenuous. The situation is further compounded by the fact that two of the partners disagree over the elements of the performance that determine its authenticity.

Summary of Focus 1 general studies

Focus 1 did not limit its research to individual case studies. Three general studies were also conducted by members of the group.

General study 03: Preserving Interactive Digital Music—The MUSTICA Initiative[50]

The first of the general studies in the arts Focus, MUSTICA, stands alone as a collaborative research project with external partners. MUSTICA was designed to fall halfway between a case study and a general study, with a short version of the case study questionnaire being administered to a large number of composers, technical assistants and other employees of two well-known French research institutes—the Institute de recherche et coordination acoustique/musique (IRCAM) and the Groupe de recherches musicales (GRM) of the Institut national de l'Audiovisual (INA). The MUSTICA researchers worked to identify the various digital components generated during artistic creation and performance processes and to confirm which ones are necessary for long-term preservation and access. They also analyzed IRCAM's long-term experience with metadata, its creation and subsequent utility in providing access to digital files. Unfortunately, language issues, geographical distances and an inability to maintain continuity among personnel involved with the MUSTICA study forced a reduction in the extent of analysis that could be undertaken on the material acquired in the interviews.

[48] Sydney Fels and Seth Dalby (2004), "InterPARES 2 Project - Case Study 15 Final Report: *Waking Dream*." Available at http://www.interpares.org/display_file.cfm?doc=ip2_c15_final_report.pdf.

[49] As noted in the case study report, because "much of the performance takes place in infra-red light, and some parts occur in the dark among the audience, no video can capture the entire performance" (Ibid., 7).

[50] See Douglas, "General Study 03 Final Report," op. cit.

General study 04: Survey of Recordkeeping Practices of Composers[51]

This general study involved the development of a Web-based survey instrument to document the record-making and recordkeeping practices of composers using digital technology. The Composers' Survey gathered information about the use of digital technology, composers' intentions and strategies for maintaining digital records, and the forms that these records might take. As such, it served, among other things, to help assess the representativeness of the recordkeeping practices associated with Canadian composer, Keith Hamel, whose electroacoustic work, *Obsessed Again...*, was investigated in case study 13.

Five hundred composers were invited to participate in the survey. With a response rate of nearly 33% of those contacted, the results show a community already familiar with digital technology, with a small minority already beginning to create works intended to be performed in a Web-based environment.

General study 07: Survey of the Recordkeeping Practices of Photographers Using Digital Technology[52]

This general study adapted the composers' survey tool to document the recordkeeping practices of photographers. In this case, no detailed case study was underway among photographers, although the widespread use of digital cameras by professionals and amateurs suggested a much larger and more mature digital practice in the field. The photography community was invited to participate in the survey through a number of online professional forums. In total, 371 photographers responded to the survey. The pool of respondents reflected all three focus groups examined by InterPARES 2: the artistic, scientific and governmental communities. Findings of the survey indicate that:

> ...professional photographers have universally embraced the transition from analogue to digital photography. The majority of photographers who responded identified their practice as completely digital, allocating the use of analogue film to the occasional personal project. Even amongst those who identified their practice as a hybrid of digital and analogue, the bulk of their images were originally created in digital form; they made only a small percentage of analogue images and most of these were eventually digitized.[53]

Findings specific to the artistic sector

The findings of the Focus 1 case and general studies indicate that there are significant differences in attitude, procedures and concerns between individual creators or small groups of artists and the business and entertainment industry. For the most part, early adopters of specific digital technologies are not part of a mainstream industry, but are instead innovators who use, adapt or are inspired by particular technological tools, be they hardware or software. This is particularly true for individual artists who explore new forms of expression through technology. This more experimental type of work also seems to rely more heavily on specific types of hardware, which often results in more problems with short-term obsolescence than is the case for those whose work is more hardware-neutral or hardware-independent.

[51] See Longton, "General Study 04 Final Report," op. cit.
[52] See Bushey and Braun, "General Study 07 Final Report," op. cit.
[53] Ibid., 29.

Records creation and maintenance practices[54]

Individuals and small organizations

There will always be artists working alone and in small groups who adopt cutting-edge technology as a tool to pursue their artistic interests or, conversely, to explore the artistic possibilities in any new technology. It is likely that attempts by archivists and other preservers to recommend that artists restrict their use of technology only to widely-implemented technologies that are subject to international standards will be perceived by many artists—especially by those who incorporate cutting-edge technology into their work—as placing an undue constraint on the artistic process. As such, it is likely that such recommendations will, to a large degree, be ignored and, hence, unproductive. Instead, the findings suggest that a more effective preservation strategy for accommodating the needs and expectations of those artists who incorporate digital technology into their work will involve securing a stronger commitment from the information technology industry for the development of technology tools with *increased interoperability*. Software applications, in particular, must provide access to widely-supported interchange formats for files. An artist may choose to work within the proprietary confines of a specific application, but should always be given the means to export the resulting works into a less restrictive technological environment.

An additional danger for artists in this group is a continuing dependency on specialized hardware, such as synthesizers[55] or customized projectors,[56] to render their work. The rapid obsolescence of hardware components can leave an artist, or subsequent performers, unable to reproduce the work as it was initially conceived and presented. Although at times difficult to implement, it is better, from a preservation stand-point, to move the definition and control of this aspect of the work from hardware to software, as is currently happening, for example, with software tools used by composers of digital musical works.[57]

Corporations and government institutions

As noted earlier, more formally established arts-related organizations are less likely to adopt experimental and cutting-edge technologies, preferring to wait for the endorsement of the marketplace before implementing new technology tools. Moreover, one of the key findings was that, in situations where technology is in rapid evolution and has failed to stabilize, corporations will often neglect or even abandon their original digital elements and focus instead on maintaining their analogue materials. This reflects the current belief among such corporations that, should the need ever arise, it will be cheaper to re-create the digital components on future digital platforms rather than attempt to perform multiple costly forward migrations through several generations of frequently incompatible technological environments.[58] Although coming from a different perspective than individual artists, this situation once again endorses the need for widespread interoperability among various file formats.

---

[54] This section provides a synopsis of some of the more salient records creation and maintenance practices in the arts based on the findings of the Focus 1 case and general studies. For more detailed discussion of these issues in relation to the Focus 1 case and general studies, see the relevant sections of the Domain 1 Task Force Report. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_2_domain1_task_force.pdf.

[55] See, for example, case study 13, *Obsessed Again...*.

[56] See, for example, case study 15, *Waking Dream*.

[57] The "resurrection" experiment in case study 13, *Obsessed Again...*, provides a good example of this strategy. See Appendix 1 in Amort, "Case Study 13 Final Report," op. cit., especially the discussion in the section titled "Towards a Strategy" at pp. 28–29. See also the discussion in the section titled "Experience with a Possible Maintenance Strategy" in the Domain 1 Task Force Report.

[58] Among the Focus 1 case studies, this view is most clearly expressed by the commercial film studio creator in case study 09(03). See Turner et al., "Case Study 09(03) Final Report," op. cit., 4 (answer to research question three).

Recordkeeping and preservation practices[59]

Individuals and small organizations

Perhaps not too surprisingly, the findings indicate that individual artists and small non-profit and for-profit entities rarely maintain recordkeeping systems that adhere to accepted practices, such as the use of standardized descriptive schemas and metadata, persistent unique identifiers for digital files, documentation of actions/transactions regarding digital file use, or the adoption of standardized work procedures. Individual artists and the small arts organizations also do not employ the services of archivists. In a small number of cases, artists have made a conscious decision that their work should not survive beyond the "life" of the technological context in which it was created due, primarily, to concerns that efforts at long-term preservation could compromise certain characteristics that the artists consider essential to the essence of their artworks, such as the ephemerality and variability of works that change each time they are reproduced. More frequently, the absence of formal recordkeeping practices reflects, in part, a lack of money or a lack of knowledge as to how to best pursue proper maintenance of digital art objects. It also reflects the fact that, because many of the artists examined have yet to suffer a catastrophic loss of some or all of the components of any of their digital works, there has been little "incentive" for these artists to begin to consider the long-term preservation implications of their decision to adopt digital technology as part of their creative processes.

This is not to suggest, however, that the artists examined have not encountered problems related to their use of technology. In fact, all of the artists admitted having had trouble finding files. In response to this issue, almost all the case study and general study respondents acknowledged that they have, with varying levels of success, attempted to implement certain procedures to help manage the digital objects that they create, such as establishing a consistent directory structure and adopting some degree of consistency in naming the objects. These procedures are most evident, and most successfully implemented, in situations where some kind of partnership requires that collaborators other than the creator have access to the material. Of particular note is the fact that a collaborative environment is also most likely to result in the creation of multiple copies of each digital object. Although unwitting, this type of activity represents the adoption of the elementary component of a preservation strategy referred to as LOCKSS (Lots of Copies Keep Stuff Safe).[60]

Corporations and government institutions

In contrast to the situation noted above for individual artists and small non-profit and for-profit entities, the findings indicate that large for-profit, non-profit and governmental organizations do utilize the services of records managers, recordkeeping systems and occasionally archivists (as was observed in two of the digital moving image case studies[61]). These larger organizations have more significant financial interests to protect in the digital assets produced and they also have the resources to implement this level of infrastructure support. Equally, the work created in these environments involves the input and contributions of many specialized individuals and departments. In this context, records management systems are often

---

[59] This section provides a synopsis of some of the more salient recordkeeping and record preservation practices in the arts based on the findings of the Focus 1 case and general studies. For more detailed discussion of these issues in relation to the Focus 1 case and general studies, see the relevant sections of the Domain 2 Task Force and Domain 3 Task Force reports. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_3_domain2_task_force.pdf and http://www.interpares.org/display_file.cfm?doc=ip2_book_part_4_domain3_task_force.pdf, respectively.

[60] LOCKSS is an international, non-profit, community initiative that provides open source software tools and support that are designed to help libraries collect, store, preserve and provide access to Web-published materials. For information on LOCKSS, see http://www.lockss.org/lockss/Home.

[61] Specifically, case studies 09(3) (Commercial Film Studio) and 09(4) (WGBH Boston).

seen as useful, if not necessary tools, to help facilitate collaboration by introducing workflow efficiencies and eliminating redundant practices. Such systems also provide a level of standardization for future access, use and re-purposing, as well as providing a level of documentation and contextualization for long-term preservation efforts.

The findings also indicate that corporate and governmental organizations tend to adopt digital technology only after it has been reasonably well-established and often for one or both of the following reasons: (1) to introduce technological efficiencies in terms of ready access to materials and (2) for the perceived financial benefits of ready access and asset re-purposing. Within these environments, the introduction and use of digital technology continues to co-exist with traditional records management practices that are already in place, such as printing to paper or accessing copies of previously recorded analogue audio or video recordings.

The "work" versus the "performance"

Regarding an issue that is unique to the performance arts creators in Focus 1, the findings indicate that there is conflicting information from creators about the need to preserve the "means of production" versus the "documentation of a performance." The first approach allows artists to re-use technological components of a work and, potentially, to continually alter, change and update the work, with or without the preservation of earlier versions. As well, some artists purposely choose to let a work "die," keeping only the documentation of each new iteration of the work. The second approach, largely favoured by archivists, curators and conservators intent on capturing the artistic output of a creator, documents the "work" through various media.

The question of what constitutes the "work" varies widely across artistic disciplines of music, performance art, experimental theatre and contemporary fine art. It ranges from the score or script to the live performance to the documentation of an installed work. Archival institutions must be clear on the scope of their acquisition mandate since this central question has implications for what will need to be preserved and how. This issue illustrates that, in the arts, *standardized* preservation criteria and procedures may not be effective or appropriate and that preservation decisions may, in many instances, need to be addressed on a case-by-case basis. This situation mirrors the experiences of museum curators and conservators working with conceptual, installation-based and performative works with digital components, who have found a wide divergence among artists in their choices of what, if anything, needs to be preserved to fully represent the artistic intent of a work.

Intellectual property rights

Some individual artists expressed little concern about establishing or protecting their intellectual property rights. This was most consistently expressed by artists working in highly experimental forms with limited commercial acceptance. Such artists are often part of small, although global, communities where their work is well-known and highly recognizable. Plagiarism in such contexts is unlikely, given the limited financial rewards and the likelihood that the deception will be uncovered and reported back to the creator quite quickly. The highly unique and idiosyncratic works of performance artist Stelarc (case study 02) offer the best example in support of this reasoning.[62]

[62] This is not to say, however, that issues related to copyright, intellectual property and patenting of technology are entirely ignored by Stelarc. In fact, as noted in the case study 02 final report, such issues "are of particular relevance to the digital entities examined" (Daniel and Payne, "Case Study 02 Final Report," op. cit., 4). Nevertheless, explicit concern for such rights appears limited mostly, if not solely, to those elements of Stelarc's works that are created by other members of his project development teams (i.e., engineers, computer scientists, visual artists, cognitive scientists, etc.), such that "copyright of the programs, codes

This "relaxed" approach to intellectual property rights exhibited by certain individual artists is obviously a major area of divergence from the corporate environment. In fact, where large corporate interests are at stake, copyright continues to be protected using traditional methods and mechanisms that have been tested and accepted by the legal system.

For individual artists, a number of current practices are, unwittingly perhaps, offering some measure of protection of their intellectual property rights. A number of individual artists indicated that although their files are usually stored on personal laptops, they also create backup copies on write-once CD-Rs to store in a second geographic location. For most artists, the practice of copying the files to external media is usually adopted to protect against loss of the original files due to theft of the laptop or failure of its hard drive, while storing the backup media in a second geographic location is seen as a means of protecting the files from small or large-scale disasters like fires and floods. However, the existence of "backup" CD-Rs containing multiple, dated copies of complete works, or versions of digital components developed during the artistic creation process, can also serve as evidence of ownership of the work.

On the other hand, although redundant storage practices can improve the chances of the long-term survival of the digital entities and, in some cases, also serve as evidence of their ownership, this practice can also create proof of ownership problems, particularly in relation to collaborative works. In fact, the ease with which artists involved in collaborative activities with other artists can exchange digital files often leads to the same files being stored on a number of different personal computers, thus making it difficult in some cases to prove exactly who originally created what.[63]

Authenticity and intent

Artistic works are very much the creative output of the artists who create them and, as such, are intrinsically connected to the lives of the artists. Implicit in this understanding is the notion that the long-term preservation of authentic digital works is, in many ways, dependent on whether the artists who created the works are still alive when the works are transferred to the custody of a designated preserver. In many cases, digital art works are based on technological innovation and invention accompanied by little, if any, recordkeeping or documentation of the techniques or practices used; information that is often critical for preservation purposes. In such cases, knowledge and understanding of the exact techniques used often disappear with the artist, thus complicating or, in some cases, thwarting the efforts of others to preserve the works following the death of the artist. In general, it is the works of individual artists and small art organizations that tend to be the least well-documented and, therefore, the most vulnerable in terms of long-term preservation.

Many artists share a common assumption about the authenticity of the materials they create. While they are alive, most artists consider themselves to be the sole arbiter of the authenticity of their digital works. When asked "whether the creator thinks that the authenticity of the digital entities is assured, and if so, why?,"[64] the majority of responses offer a surprising level of confidence that the authenticity of the digital entities is assured. For example, the performance artist Stelarc (case study 02) believes that the authenticity of the digital entities related to his

and likely the design of the technology itself belongs to the institution that designed and built the technology," while "clearly...the copyright on the interaction of body and technology, as shown in the pictures and videos of the performance, belong to Stelarc" (Ibid.).

63 For more detailed discussion of the issue of intellectual property rights in relation to the Focus 1 case and general studies, see the analysis of Domain 1 research question 3 in the Domain 1 Task Force Report.

64 Case study research question 11 (see Appendix 5).

performances is assured "primarily because of his own unique position at the centre of the entire [creation] process and the unique nature of the performance events."[65]

Although this may be true when an individual is essential to the performance of the work, the ease with which most digital materials can be duplicated suggests that the authenticity of artistic works is, in many cases, less assured than most creators seem to believe or assume. This is particularly true in situations where artists, especially those who consider themselves to be the sole arbiters of the authenticity of their works, have not documented their intent in metadata that will remain inextricably linked to the digital components of the works, since it is these metadata that will be needed by others to reproduce the works in the future following the death of the artists.

The creators of *Danube Exodus* (case study 10) and *Waking Dream* (case study 15) expressed reservations about the ongoing preservation of the authenticity of their digital works. The works associated with these two case studies involve collaborations with two or more creators. In response to case study research question no. 11—which asked whether the creators thought that the authenticity of their digital entities was assured and, if so, why—the creator in case study 10 stated that "To the extent that either the works are published, or the stewardship discussed in the above question [question no. 10] continues, yes, but beyond that, for instance beyond their own life spans, no."[66] In response to the same question, one of the co-authors (Sydney Fels) of the performance piece in case study 15 expressed concern that the digital entities "may be altered by future producers of *Waking Dream* to an extent that they will no longer be recognizable as his and his co-authors' intellectual property."[67] On the other hand, some individual artists and small collectives expressed greater concern with the preservation of their underlying artistic intent than with the preservation of the specific way in which they chose to manifest that intent in a particular work at a given time.

Regardless of the particular aspects of artistic works that artists consider to be central to the authenticity of those works, it is clear that, if there is to be any hope of preserving authentic copies of the works over the long term, additional metadata must be captured to more fully document the intent of the works from the living artists. Access to this metadata will greatly assist future long-term preservation efforts to ensure the longevity of a work beyond the life of the artist.[68]

The role of new technology

On a more optimistic note, two case studies provide early indications of how the World Wide Web is encouraging the development of a new type of recordkeeping practice that meshes with the interests of performance artists and in the dissemination of their works.

---

[65] Daniel and Payne, "Case Study 02 Final Report," op. cit., 11.
[66] Hubbard, "Case Study 10 Final Report," op. cit., 8.
[67] Fels and Dalby, "Case Study 15 Final Report," op. cit., 5. A variation of this sentiment was expressed by one of the interviewees in case study 9(2) (Digital Moving Images - National Film Board of Canada), who noted that, although the "final product is written out to film or to Betamax as a means of storage, [which] is seen as the basic means of assuring the continued existence of the animator's work in the form he approved, [it is] nonetheless...broadcast on television, or distributed in VHS or DVD form and used on uncalibrated television sets [such that] "The degradation from what the film maker saw in their head to what they're seeing at home on the tv or screen is so profound that ... some film makers just have nervous breakdowns"" (Rodger "Case Study 09(2) Final Report," op. cit., 10).
[68] For more detailed discussion of the concepts of authenticity, accuracy and reliability in relation to the arts in general, see the section titled "Conceptual Analysis: Authenticity, Accuracy and Reliability in the Literature of the Arts" in the Domain 2 Task Force Report. For further discussion of these concepts in relation to the findings of the Focus 1 case and general studies, see the section titled "Authenticity, Accuracy and Reliability in the Arts Focus Case and General Studies" in the Domain 2 Task Force Report.

Performance art and documentation is the specific focus of the Arbo Cyber, théâtre (?) study (case study 01), which introduces the concept of the *Ludosynthèse*, an interactive Web site[69] that the artists in this study created to commemorate and disseminate their past performances and also to enable users of the Web site to continue to engage, in an interactive way, with the materials of those past performances. In this case, where the performance activity took place primarily from the mid-1980s to 2001, the *Ludosynthèse* contains digital records, the sources of which are both digital and analogue materials.

The performance artist Stelarc (case study 02) offers a second example. In this case, the artist's Web site[70] is essentially an online publication that provides the general public and collaborators with access to documentation about all aspects of the artist's existing body of work, which involves performance art, dance and technology. The Web site also serves as a digital repository[71] that Stelarc relies on for his own ongoing access to materials that will serve as a basis for his future works.[72]

Focus 2 – the Sciences

Research team

The group of co-investigators conducting research on the records of scientific activities comprised experts primarily from academia with knowledge in the disciplines of Geography (Geomatics), Computer Science, Physics and Archival Science. The following is a list of researchers and research assistants who contributed to the work of the Focus 2 Task Force at some point over the duration of the InterPARES 2 Project.

Chair:

Kevin Glick	May 2003 - December 2006

Researchers:

Paul Berkman	University of California, Santa Barbara, USA—Working Group 1.2
Su-Shing Chen	University of Florida, USA—Working Group 1.2
Barbara Craig	University of Toronto, Canada—Working Group 1.2
Michèle Cloonan	Simmons College, USA—Working Group 2.2
Fynnette Eaton	U.S. National Archives and Records Administration—Working Group 2.2
Kevin Glick	Yale University, USA—Working Group 1.2
Babak Hamidzadeh	Library of Congress, USA—Working Group 2.3
P.C. Hariharan	John Hopkins University, USA—Working Group 2.2
Ken Hawkins	U.S. National Archives and Records Administration
Holger Hoos	The University of British Columbia, Canada—Working Group 2.2
Yu Lijuan	Renmin University of China—Working Group 2.2

[69] See http://www.lit.ulaval.ca/arbocyber/index_content.htm.
[70] See http://www.stelarc.va.com.au/.
[71] Interestingly, Stelarc considers the Internet to be part of his "presentation, performance and *recordkeeping system*," and views his Web site as "the central unit in that system" (Daniel and Payne, "Case Study 02 Final Report," op. cit., 7, 14, emphasis added).
[72] For more detailed discussion of the impact of technology in relation to the records management practices noted in the Focus 1 case and general studies, see the analysis of Domain 1 research question 7 in the Domain 1 Task Force Report. See also the Domain 3 Task Force Report, especially the "Findings" and "The Domain 3 Research Questions" sections.

Richard Lysakowski	Collaborative Electronic Notebook Systems Association, USA—Working Group 2.2
Reagan Moore	San Diego Super Computer Center, USA—Working Group 2.2
Eun Park	McGill University, Canada—Working Group 2.2
Xiaowei Qiu	State Archives Administration of China—Working Group 2.2
Fraser Taylor	Carleton University, Canada—Working Group 1.2
William Underwood	Georgia Tech Research Institute, USA—Working Group 2.2
Ronald Weiss	Arkival Technology Corporation, USA—Working Group 2.2
Mark Wolfe	University at Albany, State University of New York, USA—Working Group 2.2

Research Assistants:

Melissa Adams	The University of British Columbia, Canada
Bart Ballaux	The University of British Columbia, Canada
Natalie Catto	The University of British Columbia, Canada
Heather Dean	The University of British Columbia, Canada
Jennifer Douglas	The University of British Columbia, Canada
Coby Falconer	The University of British Columbia, Canada
Stephen Gage	The University of British Columbia, Canada
Erin Hanlon	The University of British Columbia, Canada
Joshua Hauck-Whealton	University at Albany, State University of New York, USA
Janine Johnstone	The University of British Columbia, Canada
Eleanor Kleiber	The University of British Columbia, Canada
Tracey Krause	The University of British Columbia, Canada
Tracey Lauriault	Carleton University, Canada
Yvonne Loiselle	The University of British Columbia, Canada
Erin O'Meara	The University of British Columbia, Canada
Christina Miller	The University of British Columbia, Canada
Carolyn Petrie	The University of British Columbia, Canada
Randy Preston	The University of British Columbia, Canada
Wendy Sokolon	The University of British Columbia, Canada
Adele Torrance	The University of British Columbia, Canada
Brian Trembath	The University of British Columbia, Canada
Melanie Wallace	The University of British Columbia, Canada
Sherry Xie	The University of British Columbia, Canada
Keum Hee Yu	The University of British Columbia, Canada

Selection of Focus 2 case studies

The scope of inquiry for Focus 2 was limited to an examination of records generated in the course of scientific activities. The Focus 2 case studies were selected for their ability to generate the data required to answer the Project's core research questions regarding the creation, management and preservation of e-science data records. They were also selected for their ability to generate the data that could be used to address research issues in the Description, Policy and Modeling Cross-domains. Because of the small membership of the Focus 2 Task Force, as compared to the other two focus groups, it was decided that the Focus 2 researchers would concentrate on conducting case studies in their current areas of research or primary knowledge, supplemented with a greater emphasis on general studies than was the case for the other two focus groups.

Summary of Focus 2 case studies

The following brief descriptions of the Focus 2 case studies identify the digital entities that were examined, while also highlighting some of the more salient findings from each study, especially with respect to the concepts of authenticity, reliability and accuracy as these are understood in the sciences. For a more detailed contextual analysis of these case studies in relation to the records creators and the nature of the activities resulting in document creation, see the "Characterization of the Case Studies" section in the Domain 1 Task Force Report. See Appendix 3 for a complete list of participants responsible for each case study in this focus.

Case study 06: CyberCartographic Atlas of Antarctica[73]

The focus of this case study is the Cybercartographic Atlas of Antarctica (CAA), an online atlas that portrays, explores and communicates the complexities of the Antarctic continent for education, research and policy purposes. The Atlas is the key deliverable of a four-year (2003-2007) research project that was led by the Geomatics and Cartographic Research Centre (GCRC), an organized research unit in the Department of Geography and Environmental Studies at Carleton University, Ottawa, Canada.

The CAA is a complex system that is open source, interoperable, renders distributed data, includes scientific data visualization, is multisensory and incorporates multimedia data in a very wide array of formats from a number of international sources. In collaboration with experts from a number of disciplines—including Cartography, Geography, Psychology, International Studies, English, Cultural Mediation, Music Studies, Industrial Design and Computer Science—these data are used to develop theme-specific modules for use by the general public to facilitate knowledge sharing in multi-disciplinary science.

It was not the intention of the CAA Project to generate or collect substantive new data but to bring together selected existing datasets in a new multimedia form including experimental work with virtual reality and render these in a dynamic and engaging fashion. Distributed data are rendered "on the fly" in maps, charts, tables and text.

The CAA's design incorporates recent specifications proposed by the Open Geospatial Consortium (OGC) and the International Organization for Standardization (ISO) that are enabling the development of the "spatial Web." This spatial Web is making it possible for users to easily find, access and process digital geospatial data over the Internet. It is, of course, also creating new long-term preservation challenges.

The creators in this study rely on various processes and procedures to ensure the quality of the data in the CAA, including the professional practices and authority of the institutions from which the data are derived, implementation of access and security controls, and the use of peer review together with adherence to cartographic professional practices to choose the right level of data accuracy and to select cartographers for the right representation. These processes and procedures are very much reliant on metadata and professional practices.[74] Regarding authenticity, this is assessed by the creators of the CAA through analysis of what in Geography and Geomatics is typically referred to as "data lineage." "Lineage," which is one of at least seven

[73] See Lauriault and Hackett, "Case Study 06 Final Report," op. cit.
[74] Ibid., 14–15, 25–27, appendices P, T.

elements comprising "spatial data quality" and is a mandatory metadata element for any digital object that forms part of the CAA,[75] captures information about the chain of transmission of a dataset—from the moment the data in the dataset were originally recorded—that brought the dataset to the user. In other words, data lineage speaks to the *history* of a dataset; that is, its lifecycle from data collection to its many stages of compilations, corrections, conversions and transformations, including the generation of new interpreted products.[76]

Although clearly not a study of a recordkeeping or a preservation system, this case study raises the important and timely issue of how copies of digital records and related objects that an organization makes available over the Internet can be preserved authentic. As the findings of this study suggest, a conscious commitment to, and consistent reliance on, open source software, especially open source software that is highly interoperable, supported by a robust repository system that incorporates a classification system and "a multimedia metadata schema that meets both archival requirements and geospatial standards,"[77] is likely one of the most effective strategies for helping ensure the long-term sustainability and preservation of complex digital entities like the CAA.

Case study 08: Scientific Data Records from a NASA Spacecraft Mission[78]

The focus of the case study is the data records of the Mars Global Surveyor (MGS) mission and the Planetary Data System (PDS). This case study was selected because it was also designed to collect data used in two general studies, general study 12 (Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data) and general study 06 (A Bayesian Belief Network: Supporting the Assessment of the Degree of Belief that a Recordkeeping System Maintains Authentic Digital Records).

Launched in 1996, the MGS was the first mission of the National Aeronautics and Space Administration (NASA) Mars Surveyor Program. The MGS spacecraft contains instruments that send observational data back to the Jet Propulsion Laboratory Spacecraft Control Center where they are stored as experimental data records in a Project Database. Spacecraft Operations Planning Computers at the Principal Investigator's site and at seven discipline nodes—Atmospheres, Geosciences, Imaging, Navigation Ancillary Information Facility, Plasma Physics, Rings and Small Bodies—are used to create reduced data records.

The designers of the PDS adopted the use of self-describing data files—that is, files that describe the file format of the attached data as well as the context in which the data were created—as a preservation strategy. The use of self-describing files eliminates the need for converting scientific data to other formats to maintain their accessibility over time, which is one of the key elements supporting an experimental archival preservation strategy called Collection-based Persistent Object Preservation (POP).[79] Using this strategy, technology obsolescence requires only the migration of an interpreter, a data access library and a viewer, thereby

---

[75] Ibid., 14, 26.

[76] As is discussed in more detail in the Domain 2 Task Force Report, data lineage can also be characterized as "data provenance" and, thus, roughly analogous to an amalgamation of the archival concepts of provenance and chain of custody.

[77] Lauriault and Hackett, "Case Study 06 Final Report," op. cit., 29.

[78] See William Underwood (2005), "InterPARES 2 Project - Case Study 08 Final Report: Mars Global Surveyor Data Records in the Planetary Data System." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs08_final_report.pdf.

[79] For a brief description of POP within the context of this case study, see Underwood, "Case Study 08 final Report," op. cit., 34. For a more detailed discussion, see Reagan Moore et al. (2000), "Collection-Based Persistent Digital Archives - Part 1," *D-Lib Magazine* 6(3). Available at http://www.dlib.org/dlib/march00/moore/03moore-pt1.html; and Reagan Moore et al. (2000), "Collection-Based Persistent Digital Archives - Part 2," *D-Lib Magazine* 6(4). Available at http://www.dlib.org/dlib/april00/moore/04moore-pt2.html.

circumventing the need to continuously update or create new software for each of the file structures. The fact that the PDS has been operational since 1989, during which time it has not been necessary to update (convert or migrate) any of the self-described data products to other data formats, seems promising.

Concerns regarding the reliability and accuracy of the digital entities in this study are addressed in several ways. First, the creators in this study make use of various data processing plans, manuals, specifications and workbooks to guide processing, preparation and transfer of satellite data slated for long-term preservation in the Planetary Data System (PDS). Second, the data are subjected to a peer review process that is designed to check the digital entities for accuracy, reliability, suitability for archiving and to ensure that the PDS data processing, cataloguing and transmission standards have been followed. Finally, all data destined for the PDS are parsed by a computer program that validates the format and content of the "product label" metadata as well as the integrity of each data file using checksums.[80] Regarding authenticity, it is noted that "Project team members, PDS managers and engineers and other Planetary Scientists do not traditionally use the term authentic to characterize the data products that they create, maintain and use. They are concerned that the data records are complete, reliable, accurate, and that the integrity of the data record is assured."[81] However, as the case study report further clarifies: "Given the definition of authentic digital record as "a digital record that is what it purports to be and is free from tampering or corruption," one can conclude that, due to the emphasis on completeness and reliability of the planetary science data records, the peer review, role-based authentication of access to archived data products and data integrity checks, the scientific data records are maintained authentic. However, it is unlikely that the creators or maintainers of the records would customarily use that term to characterize their quality."

Case study 14: Archaeological Records in a Geographical Information System[82]

This case study investigates archaeologists' attitudes towards recordkeeping, especially in terms of what they consider authentic and reliable records and the value of these concepts within archaeology. The immediate goal of this case study was to gain insight into the creation processes of Geographic Information System (GIS)-related data and records created and used in the course of archaeological research, with an emphasis on better understanding how archaeologists themselves view and, subsequently treat, their data and the records resulting from the use of those data. The more general objectives of this study were to help answer questions about the nature of digital archaeological records in general; about how the increased reliance on GIS-supported research is impacting the archaeological community; and, especially, about how both of these issues are playing out in the recordkeeping habits of archaeologists.

Ultimately, this research was conducted with the objective of facilitating the development of an appropriate records management framework in which archaeologists and other scientists can approach the long-term preservation of the complex datasets and outputs that are commonly associated with a GIS. The long-term preservation of such datasets and outputs has become a critical issue with regard to enabling the kind of multidisciplinary research crucial to modern scientific knowledge.

[80] See Underwood, "Case Study 08 Final Report," op. cit., 10–11, 23–24.
[81] Ibid., 24.
[82] See Richard Pearce-Moses, Erin O'Meara and Randy Preston (2004), "InterPARES 2 Project - Case Study 14 Final Report: Archaeological Records in a Geographical Information System: Research in the American Southwest." Available at http://www.interpares.org/display_file.cfm?doc=ip2_c14_final_report.pdf.

The case study investigates the recordkeeping activities of those individuals responsible for the creation, maintenance and use of the Coalescent Communities Database and GIS that was created for the Center for Desert Archaeology in Tucson, Arizona, in support of a two-year U.S. National Science Foundation grant-funded research project. The database and related GIS consist of compilations of pre-recorded archaeological site data from multiple sources. These datasets exist in various formats, such as paper, spreadsheets and pre-existing databases.

As is true of virtually all applications of a GIS, the central function of the GIS in this case study is information management. In this particular instance, however, the process of incorporating large legacy datasets—many of which were collected long ago (more than 100 years ago in some cases) and originally recorded in analogue form and subsequently digitized—from diverse sources into a single database raises serious concerns about establishing and maintaining the authenticity, accuracy and reliability of the data and, hence, the records generated by the GIS from the data in the consolidated database.

Regarding the datasets used to populate the GIS database in this study, it is noted that an assessment of the accuracy and quality of each dataset may involve various processes or procedures. These may include preliminary "audits" of datasets both before and after they are ingested into the database; "routine" checks of the data in the database for data redundancies, errors and omissions by a volunteer who is a retired archaeology professor; and "spot" checks by the creator of data entered into the database by the volunteer to ensure that the data are being entered accurately. However, since there are no formal, documented data assessment rules, procedures or guidelines that control these activities,[83] the overall process is best characterized as idiosyncratic and ad hoc. The relative laxness of the data assessment approach used by the creator in this study is based, to a large degree, on his assumption that the reliability and accuracy of the data used to populate the database is determined by the reliability of the data *source*. In other words, the creator believes that if the source of the data can be assumed to be trustworthy, then the data acquired from that source can be assumed to be reliable and accurate. Moreover, this same argument also serves as an implicit means of "establishing" the authenticity of the data in that the creator assumes the data are authentic on the grounds that they are acquired from an official state repository and from fellow professional researchers whom the creator trusts to collect and maintain their research data using standard, professional archaeological procedures.[84]

To help assess the representativeness of the findings from this case study, Focus 2 researchers later conducted general study 09 (Survey of the Digital Recordkeeping Practices of GIS Archaeologists Worldwide), which is summarized later in this chapter.

Case study 19: Preservation and Authentication of Electronic Engineering and Manufacturing Records[85]

This case study reports on an engineering experiment to develop an open-source preservation format for digital computer-aided design (CAD) records of solid models. The domain of the case study is science-based manufacturing of complex, high-assurance, high tolerance machined piece parts for the U.S. government. The business owner has an ongoing need to access and use the preserved records for business purposes over a long period of time (50+ years) with the assurance that the records remain accurate, reliable and authentic.

[83] See Pearse-Moses et al., "Case Study 14 Final Report," op. cit., 7–8.
[84] Ibid., 2, 15, 24–26.
[85] See Kenneth Hawkins (2006), "InterPARES 2 Project - Case Study 19 Final Report: Preservation and Authentication of Electronic Engineering and Manufacturing Records." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs19_final_report.pdf.

The records involved contain information about complex geometric and topological measurements and relationships of various parts of three-dimensional objects. The case study involves analysis of an experiment that involved: (1) abstraction of this information from proprietary CAD formats; (2) expression of this information into enhanced logical forms that support reasoning about part shape and manufacturing actions; (3) rendering of the proprietary CAD formats into an open-source archival format; (4) transmission of the records in open-source format across a trusted network for ingestion into a persistent archives; and (5) retrieval of the archived records for verification of their continued authenticity, reliability and usability.

The intent of the experiment was to preserve not only the geometric specifications of the models but also their semantically encoded metadata, joined to make a "new logical preservation format" for archival purposes. By "logical preservation format," the experiment partners in this case study meant a format encompassing not only the fixed form and content of information representing the models, but also instructions encoded within their metadata so that reasoning engines of the future can conduct "proofs" against the objects to authenticate them as fit to support the procedural action for which they were designed to be used.

Concerns regarding the authenticity, reliability and accuracy of the digital entities in this study are addressed in several ways. First, accuracy and reliability are in the hands of the designer and his/her adherence to design modeling standards such as the ANSI Y-14.5 tolerance standard, which provides for the assurance of measurements down to the millionth of an inch. Second, the CAD system used to create the initial CAD entities integrates rigorous model checks that can be used alert users whenever they attempt to create bad geometry. Finally, a quality guide is used to help control the digital entity production and manipulation processes. However, because of the extreme level of accuracy achieved by these processes—especially in relation to the "enhanced semantic knowledge about the geometric and topologic characteristics of the solid models"—the resulting design records cannot be accurately translated into the final archival format, at least not within the context of the engineering experiment.[86] The results of the engineering experiment show that, by the exacting standards of the engineering experiment, the final archival entities are unable to preserve all the advanced information contained in the "knowledge-enhanced" entities produced from the original CAD entities and, as such, "cannot be said to be authentic to a satisfactory degree." In short, the experiment in this case study demonstrates that it is not possible, at present, "to preserve all elements and components of the entities in a persistent archives."[87]

Case study 26: MOST Satellite Mission - Preservation of Space Telescope Data

The focus of this case study is the preservation of astronomical data from the MOST (Microvariability & Oscillations of STars) Satellite Mission, Canada's first space telescope. MOST is designed to monitor variations in the brightness of stars with unprecedented precision and time coverage. Launched in late June 2003, MOST is funded by the Canadian Space Agency (CSA) and operated jointly by Dynacon Inc., the University of Toronto Institute for Aerospace Studies (UTIAS) and the University of British Columbia (UBC).

This case study examines how the MOST researchers manage the data that are transmitted from the space telescope. In particular, the case study investigates how technological restraints complicate the creation, maintenance and preservation of the MOST satellite data and how future researchers can be assured that the preserved data are reliable, authentic and accurate. In so

[86] Hawkins, "Case Study 19 Final Report," op. cit., 11, 16.
[87] Ibid., 17.

doing, this study explores the effect of the business process on the various phases of the life-cycle of the data and records. The study also examines how researchers interpret the concepts of accuracy, reliability and authenticity in the context of scientific research.

Perhaps not too surprisingly, similar procedures to those noted earlier for the NASA Spacecraft Mission (case study 08) are used to ensure the authenticity, reliability and accuracy of the digital entities in this study. The creators in this study have implemented a rigorous and systematic, two-stage data integrity process. The first stage involves an automated, computer-executed, checksum analysis. The second stage involves a hands-on "intellectual analysis" performed by the project's instrument scientist. As well, some data are processed through a model that, among other things, helps the researchers spot and isolate any inaccuracies in the data that might require that the data be reprocessed.[88] Regarding authenticity, the report notes that "there is no reason for the creator to assume that the digital entities are not authentic." This level of confidence is not specified is likely a consequence of the controlled and secure nature of the sources of the data used in this study,[89] together with the fact that there is only one researcher who is responsible for the creation, maintenance and preservation of the raw data files received from the space telescope and gathered from other sources, and one researcher who is responsible for processing (i.e., "reducing") the data files during the interpretation phase of the research.[90]

Summary of Focus 2 general studies

General study 01: Persistent Archives Based on Data Grids[91]

This study focuses on the San Diego Supercomputer Centre's project to develop a prototype for a "persistent archive"[92] based upon Data Grid technology for the U.S. National Archives and Records Administration (NARA). The study examines the minimal capabilities needed within grid technology for preservation of governmental records, focusing on activities related to the preservation of NARA's selected digital holdings.

As discussed in this study, a Data Grid is

> the set of abstractions that manage differences across storage repositories, information repositories, knowledge repositories and execution systems. Data Grids also provide abstraction mechanisms for interacting with the objects that are manipulated within the grid, including digital entities (logical namespace), processes (service characterizations or application specifications) and interaction environments (portals). The Data Grid approach can be defined as a set of services and the associated application programming interfaces (APIs) and protocols used to implement the services. The Data Grid is augmented with portals that are used to assemble integrated work environments to support specific applications or disciplines.[93]

---

[88] See Ballaux, "Case Study 26 Final Report," op. cit., 11–12.
[89] In addition to the data received from a single space telescope, whose data collection functions are under the sole control of select project researchers, certain other data are received from a few other sources that are assumed to provide authentic and reliable data, such as the orbital information downloaded from the North American Aerospace Defense Command (NORAD) Web site (Ibid., 7).
[90] Ibid., 5, 13.
[91] See Moore, "General Study 01 Final Report," op. cit.
[92] Persistent archives are "collections of digital entities that map from unique attribute values to a global, persistent identifier," (Ibid., 5) which are distinct from "persistent storage systems ", where the emphasis is on providing "archival media that have a very long shelf life, such as heavy-ion beam encoded disk, film, etc." (Ibid., 6).
[93] Ibid. For more information about Data Grids, see Fran Berman, Geoffrey Fox and Tony Hey (eds.), *Grid Computing: Making the Global Infrastructure a Reality* (Chichester, West Sussex: John Wiley & Sons, 2003).

Preservation environments for digital records are successful when they can separate the digital record from any dependence on the original creating infrastructure. Data Grid technology, which supports the management of distributed records, provides the software needed for infrastructure independence. Operating under the assumption that the capabilities of virtual data grids can be used to implement the traditional archival processes of appraisal, accession, arrangement, description, preservation and access, this study maps Data Grid capabilities to the archival processes required by the InterPARES 1 Model of the Preservation Function.[94]

General study 06: A Bayesian Belief Network: Supporting the Assessment of the Degree of Belief that a Recordkeeping System Maintains Authentic Digital Records[95]

One of the research results of InterPARES 1 was the Authenticity Task Force's (ATF's) formulation of the *Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records*.[96] Quoting the ATF report: "A presumption of authenticity will be based upon the number of requirements that have been met and the degree to which each has been met. The requirements are, therefore, cumulative: the higher the number of satisfied requirements and the greater the degree to which an individual requirement has been satisfied, the stronger the presumption of authenticity." The purpose of this general study is to investigate a method of assessment for determining just how these requirements can actually be applied to assess the authenticity of a creator's digital records.

The approach used in the study is based on an interpretation of probability suggested by Bayesian theory, which holds that the concept of probability can be defined as the degree to which a person believes a proposition. Bayesian theory also suggests that Bayes' theorem can be used as a rule to infer or update the degree of belief in light of new information.

The researcher in this study conducted an experiment using a Bayesian Belief Network (BBN)[97] and data from case study 08 (Mars Global Surveyor Data Records in the Planetary Data System). A person, using data from the case study as evidence, was asked to express his degree of belief in the truth or falsity of the requirements at the terminal nodes of the BBN. As the degrees of belief were entered, the beliefs were propagated to the target hypothesis: "The Recordkeeping System is trusted and maintains authentic digital records." The result of this particular experiment was a very high degree of belief in the hypothesis.

Additional refinement and experiments are needed before the Bayesian approach outlined in this study can be considered a reliable and practical tool supporting the assessment of authenticity. However, among the results of this investigation is the identification of a number of conditional dependencies among the variables (requirements, attributes in record schema) that are not explicitly represented in the original statement of requirements. Additional consideration

---

[94] See Preservation Task Force (2002), "Appendix 5: A Model of the Preservation Function," version 6.0, in Duranti, *Long-term Preservation*, op. cit., 253–292. Online reprint available at http://www.interpares.org/display_file.cfm?doc=ip1_ptf_model.pdf. The InterPARES 1 Model of the Preservation Function was used in this general study because, at the time the study was conducted, the InterPARES 2 Chain of Preservation Model had not yet been developed.

[95] See Underwood and Isbell, "General Study 06 Final Report," op. cit.

[96] The *benchmark requirements* are the core information about a creator's digital records, and about the procedural controls exercised by the creator over the creation, handling and maintenance of its records, that set forth a basis for presuming or verifying the authenticity of the creator's digital records. Related to these are the *baseline requirements*, which are the minimum conditions necessary to enable the preserver to attest to the authenticity of copies of the creator's inactive digital records. For more information, see Authenticity Task Force, "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records" in Duranti, *Long-term Preservation*, op. cit., 204–219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf. An abridged version of the benchmark requirements is provided in Appendix 21a.

[97] A BBN is a graphical notation with an associated set of probability tables.

needs to be given to the relative contribution of requirements to the overall assessment of the authenticity of the records maintained in the recordkeeping system. Furthermore, there seem to be additional requirements (variables), particularly in the security area, that affect the authenticity of the maintained records but that were not included in this experiment.

General study 09: Survey of the Digital Recordkeeping Practices of GIS Archaeologists Worldwide[98]

This general study summarizes the results of a forty-question, Web-based survey of the digital recordkeeping practices of GIS archaeologists worldwide. Invitations to participate in the survey were e-mailed to nearly 900 GIS archaeologists from sixty-nine countries worldwide. Additional invitations were posted to various professional archaeology- and GIS-related listservs and Internet discussion groups. The survey was available online for thirty days, during which 157 complete, or nearly complete, surveys were submitted.

Reliance on GIS technology has increased dramatically in archaeological research in the past two decades to the point where it is quickly becoming the de facto tool for facilitating data amalgamation, manipulation, synthesis, analysis and preservation in archaeological research. This fact raises concerns regarding the impact of GIS-facilitated research on archaeologists' ideas and attitudes towards digital recordkeeping practices, particularly in relation to their views towards creating and trusting digital records. The goal of the survey, therefore, was to gather and analyze baseline data about the existing digital recordkeeping knowledge and practices of GIS archaeologists worldwide to help gauge the current level of awareness and understanding within the global archaeological community about: (1) digital preservation issues, (2) digital recordkeeping practices and (3) the potential impact of such practices on the long-term preservation of accurate, reliable, authentic and accessible digital archaeological data and research records. This study was designed to complement the case study 14 (Archaeological Records in a Geographical Information System) research by providing more comprehensive and broad-based data on the recordkeeping practices of GIS archaeologists that could be used to help assess the representativeness of the findings in the case study.

The survey findings indicate that, despite the general perception of archaeologists as being meticulous and thorough in their documentation and recordkeeping practices, many GIS archaeologists currently engage in idiosyncratic or ad hoc file creation, management, preservation and/or documentation practices that have the potential to seriously compromise the accuracy, reliability, authenticity and accessibility of the files they create, especially over the long-term. One possible explanation for this apparent incongruity is that, until relatively recently, archaeologists typically have had little or no formal GIS training and have, instead, learned their GIS practices through a self-taught trial-and-error approach. These findings concur with the case study 14 findings.

Nevertheless, to the extent that the results of the survey can be considered representative of the habits of GIS archaeologists in general, it appears that, on the whole, there is a considerable and growing level of awareness among GIS archaeologists of the many technical, administrative, professional and societal issues surrounding the long-term preservation of their archaeological GIS data and research records.

[98] See Preston, "General Study 09 Final Report," op. cit.

General study 10: Preservation Practices of Scientific Data Portals[99]

Geomatics and other science data are increasingly being discovered and accessed in data portals. Portals have a variety of names, such as data repositories, clearinghouses, catalogues, archives, geolibraries and directories. In this study, the term portal is used to encompass all of these. Portals can provide all or some of the following services: data search and retrieval, item descriptions, display services, data processing, platforms to share models and simulations, and data collection and maintenance. Much, but not all, of the data derived from portals are "raw" or unprocessed and require the user to interpret, analyze and/or manipulate them.

This general study was undertaken to collect structured information about the actual practices, standards and protocols currently used by existing science data portals in an attempt to understand how these types of entities address and demonstrate the lineage, reliability and authenticity of their data assemblages. To this end, a semi-structured survey of a random collection of thirty-two science data portals was conducted. The portals surveyed represented many different communities of practice in the sciences such as health, astronomy, biology, engineering, statistics, genetics, geosciences and ecology, to name a few.

Since the survey was undertaken for exploratory purposes, the sample size from each scientific discipline is small, thus limiting cross-disciplinary analysis. The study does, however, provide a deeper understanding of preservation practices in the natural and physical sciences in relation to data scientific portals. The report also includes an exploratory literature review that considers the importance of issues such as accuracy, reliability and authenticity in the management of scientific data exchanged through portals.

Among other things, this study demonstrates the heterogeneity of science data, the attention paid to scientific data quality, the complexity of the scientific methods that create and utilize these data and the myriad types and forms of scientific data that are created. In so doing, this study highlights the fact that the reliable and accurate creation of digital scientific records, and their authentic maintenance and preservation over the long term is a complex issue that needs to be addressed on several levels, This, in turn, emphasizes the need for a closer relationship between scientists and preservers to ensure that the digital records that scientists create and deem worthy of retention are created and managed in a manner that is conducive to their long-term preservation, and to ensure that preservers are able to preserve and provide access to the records in a manner that is sensitive to the often unique and varied requirements related to both the preservation and use of these types of records.

General study 12: Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data[100]

This general study presents the results of a walkthrough of an early draft of the InterPARES 2 Project Chain of Preservation (COP) model using case study data. Walkthroughs using case study data are an effective way to test whether a model, design, program code or user interface achieve what is intended and to improve the quality of the product. A walkthrough is a peer group review of any information system product. A walkthrough of an activity model, such as the COP model, is concerned with the functionality of the system.

---

[99] See Lauriault and Craig, "General Study 10 Final Report," op. cit. Additional findings and discussion stemming from this study can be found in Tracey P. Lauriault et al. (2007), "Today's Data are Part of Tomorrow's Research: Archival Issues in the Sciences," *Archivaria* 64 (Fall): 123–179.

[100] William Underwood, Kevin Glick and Mark Wolfe (2007), "InterPARES 2 Project - General Study 12 Final Report: Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs12_final_report.pdf.

The COP is a generic model of the processes involved in creating, maintaining, selecting and preserving authentic digital records. The COP model is prescriptive as well as descriptive. It prescribes criteria for determining whether digital records can be presumed to be authentic and it prescribes a method for applying these criteria. However, because the model was developed via functional decomposition rather than generalized from specific archival objects and information used to manage those objects, the following question arises: How can a user be sure that the model actually applies to the creation, maintenance, selection and preservation of digital records? The purpose of this study, therefore, is to demonstrate whether the COP model applies to specific record-making, recordkeeping and preservation systems for digital records and to refine and validate the model by conducting a walkthrough of the model using InterPARES 2 case study 08 (Scientific Data Records from a NASA Spacecraft Mission) data. As noted earlier, case study 08 was designed to collect information about the Planetary Data System (PDS) Space Science Data Archive.

NASA refers to the PDS as an active archive. Copies of the scientific datasets are transferred to the National Space Science Data Center (NSSDC) for long-term preservation. The NSSDC is referred to as a deep archives. The PDS activities of data preparation and management of datasets in the PDS Archive are similar to the COP activities “A2. Manage Records in a Record-making System”[101] and “A3. Manage Records in a Recordkeeping System.” The management at the NSSDC of scientific datasets from the PDS and from other space science disciplines appears to be similar to the COP activity “Select and Preserve Records.” However, further analysis of case study 08, together with the results of the walkthrough using the case study’s data, reveals that selection (appraisal) and preservation were central aspects of the PDS design and operation.

An interesting aspect of the walkthrough using the case study 08 data is that the description activity seems to take place during the record-making phase rather than during the preservation phase following the transfer of the records to the organization responsible for the long-term preservation of the records, as is prescribed in the COP model. Another interesting aspect of this case study is that parts of the appraisal and validation activities take place in the PDS *before* the recordkeeping phase, while in the COP model they take place *after* the recordkeeping phase. A possible explanation for this relates to the PDS management decision to actively involve scientists in the archiving process, which is based on the rationale that the scientists who create and use the data products are better able to describe and appraise them than are archivists (or scientists) far removed from the mission and data creation. Because of the expense of space science missions, the investment dictates early description, appraisal and validation of the datasets.

Of the sixty-eight lowest-level activities in the draft COP model, data from the case study were found to correspond to forty-six of those activities. There is no corresponding data in the case study for seven of the COP activities, while no data were collected for fifteen of the NSSDC activities that would correspond to the long-term preservation activities of the COP model.

These findings do not invalidate the COP model, rather they emphasize that it is a more general model of recordkeeping activities than the activities of scientific data recordkeeping in certain domains. In fact, the walkthrough of the COP model using the case study 08 data demonstrates that, overall, there is an interpretation of the record-making, recordkeeping and some of the preservation activities of the COP model in the domain of scientific data archives. That is to say, the COP model is satisfiable in this domain. A more thorough validation of the COP model would require further walkthroughs of the COP preservation activities using case study data with regard to the NSSDC preservation activities and walkthroughs using case study data from other scientific domains.

[101] This activity was called “Manage Records Creation” in the earlier draft version of the COP model used for the walkthrough.

Findings specific to the scientific sector

The findings of the Focus 2 case and general studies indicate that, while there are certain obvious commonalities in the record-making, recordkeeping and record preservation concerns among creators in the sciences, there are also some significant differences in attitude, procedures and concerns between scientists in different disciplines, as well as among scientists within the same discipline in some cases. For the most part, the main focus in the sciences is on *data*, while the concept of record, at least in the archival sense of the term, receives far less attention or consideration. To some degree, the ambivalence of many scientists toward the archival concept of record is due to the fact that the concept of record as used and understood by most scientists is imbued with very specific and particular meaning that, for the most part, stands in stark contrast to the way the term is used and understood by archivists and other preservers. A somewhat similar situation is evident in relation to the concepts of authenticity, reliability and accuracy. Finally, despite the inherent experimental nature of scientific inquiry, evidence for the adoption and use among scientists of innovative, new or cutting-edge digital technologies is mixed. One indication of this is the use of customized and/or custom-built software applications, which was actually found to be higher among Focus 3 (government) creators than among Focus 2 creators. In fact, *all* of the creators investigated in the Focus 3 case studies were found to use substantially customized, commercial-off-the-shelf (COTS) software applications and/or custom-built applications (as well as unmodified COTS applications in most cases), while the same was not the case among the Focus 2 creators. One likely explanation for this is that, because scientists need to be able to easily share and exchange data, analysis results, etc., their activities are likely more sensitive to and influenced by the push to establish and follow industry standards regarding data creation and maintenance, as well as by the need to rely on applications and data formats with more robust interoperability functionalities and capabilities.[102]

The concept of record

As was noted for the arts creators examined in Focus 1, the fundamental concept of what constitutes a "record" in a scientific context is open to wide and varied interpretation. This fact was made evident to Focus 2 Task Force members during both the review of the science literature and during the case and general studies research. The MIT Appraisal Project serves as a good example of the former. On the one hand, the MIT project, which focused on the nature of scientific activity and the scientific record, defines the scientific record as including experimental designs, documentation of instrumentation, experimental data records and analyses of experimental results; all entities that are in close alignment with the way the term record is interpreted in most archival contexts. On the other hand, the project also includes in its definition of the scientific record the *publication* of results in technical reports, conference proceedings and journal articles; all of which, because they are publications, are not considered by most archivists to be records—except in very specific and limited contexts (such as where the offprint of an author's journal article is retained by the author as an evidentiary record of the act of publishing the article).[103] The process of publication was also seen by some of the creators in Focus 2 as

[102] For more detailed discussion of the impact of technology in relation to the records management practices noted in the Focus 2 case and general studies, see the analysis of Domain 1 research question 7 in the Domain 1 Task Force Report. See also the Domain 3 Task Force Report, especially the "Findings" and "The Domain 3 Research Questions" sections.

[103] See Joan K. Hass, Helen W. Samuels and Barbara T. Simmons, *Appraising the Records of Modern Science and Technology: A Guide* (Cambridge, MA: Massachusetts Institute of Technology, 1985); and Joan K. Hass, Helen W. Samuels and Barbara T. Simmons (1986), "The MIT Appraisal Project and its Broader Implications," *American Archivist* 49(3): 310–314.

instrumental in conferring legitimacy or "recordness" to the findings and outputs of their research.

Further complicating the situation in the sciences is the fact that, because of the inherently experimental and probatory nature of scientific inquiry, scientists view their research data, outputs and findings as *provisional* and continuously subject to revision by themselves and/or by other scientists. At a very fundamental level, it appears that this sense of the inherent impermanence of scientific research permeates and may, in many cases, even undermine the importance or relevance that scientists place on the characteristics of fixed form and stable content, both of which are central to the concept of recordness in an archival context. If so, then it is not too difficult to surmise how this attitude might shape, at a theoretical level, the basic concept of what constitutes a record in the eyes of many scientists, as well as, at a more practical level, how scientists perceive and carry out the creation, management, use and disposition of science records.[104]

Authenticity, reliability and accuracy

Of the three concepts of authenticity, reliability and accuracy, it is the concept of accuracy, especially in relation to what is often collectively referred to as "data quality," that receives the greatest attention in the science literature and in the Focus 2 case and general studies. This appears to be due in part to the fact that scientists typically focus more on data than on records, and in part to the centrality of data accuracy or quality to the validity of scientific research. Moreover, as the Focus 2 studies show, these concepts are interpreted and addressed in different ways by creators both between science disciplines and even, in some instances, within like disciplines; not unlike the situation noted earlier with creators in the arts. As the results of the Science Data Portals study clearly demonstrate, this situation is further compounded by the vast scope and breadth of scientific research, and the myriad forms and types of data records that this situation produces. To begin to successfully untangle and address the complex issues surrounding the reliable and accurate creation of digital scientific records and their authentic maintenance and preservation over the long term, preservers must, as a first step, develop a clearer understanding of the many nuanced ways that the fundamental archival concepts of authenticity, reliability and accuracy play out in the sciences.

Observations derived from the Focus 2 studies suggest that accuracy, which tends to be the most important data/records-related issue for scientists, is most commonly associated with, and addressed in relation to, the risk of having inaccurate data. Not surprisingly, it is also evident that the more critical the need for accuracy is, the more rigorous are the quality requirements and checks that tend to be established. Likewise, the more automated the data creation process is, the more technical the quality assessment procedures tend to be and, consequently, the more reliant the creators are on the technical systems in place, with human checks typically playing a secondary role in the assessment process.[105] On the other hand, in situations where very high data quality and accuracy levels are not as critical, attainable or practical, and where the data creation and quality assessment processes are less automated and more reliant on human

[104] For more detailed discussion of these findings on the concept of record in relation to the Focus 2 case and general studies, see the analysis of Domain 1 research question 4 in the Domain 1 Task Force Report.

[105] See, for example, case studies 08 (NASA Spacecraft Mission), 19 (Electronic Engineering and Manufacturing Records) and 26 (MOST Satellite Mission).

intervention, concerns about professional practice and reliance on the trust associated with the integrity and authority of external data providers are given greater emphasis by the creators.[106]

The Focus 2 studies also reveal that, although, in many cases, scientists are unlikely to use the terms "authentic" or "authenticity" in relation to the data and records they generate or use in their research, the fundamental archival concepts underlying these terms are, nevertheless, often addressed. In some situations this is done implicitly, such as is noted in the Archaeology GIS and MOST Satellite Mission studies, where the authenticity of the datasets in both studies is assumed based on a presumption that the data are obtained from trustworthy sources. In other instances, the assessment of authenticity is more explicit, even if what is being assessed is not always identified using the terms authentic or authenticity. The Cybercartographic Atlas and NASA Spacecraft Mission studies provide good examples of this more explicit approach to authenticity issues. In the former study, the concept of authenticity (or at least its identity component[107]) is imbedded as an element (i.e., data lineage) within the more general concept of spatial data quality.[108] In the latter study, the concept of integrity—which, as noted, is one of the key components of authenticity—is explicitly identified as one of the characteristics of data records that is of primary concern to planetary scientists.[109]

Science data archives

Scientists in many fields recognize the value of archiving digital scientific data, distributing the data to scientists in the field and preserving the data for future use. In fact, as the Scientific Data Portals study (general study 10) clearly demonstrates, there are scientific data "archives" in many fields of science, particularly in the areas where there is substantial government sponsorship; for example, the Space Sciences and High-Energy Physics.[110] Some of the scientific data archives include digital copies of scientific publications related to the scientific data. However, most institutional technical reports, conference records (proceedings) and journal

[106] See, for example, case studies 06 (Cybercartographic Atlas of Antarctica) and 14 (Archaeological Records in a GIS). It is noted, however, that the creator of the Atlas, unlike the creator of the Archaeological GIS, also relies heavily on good metadata provided by external data sources and on the metadata related to the Atlas modules themselves.

[107] As clarified by the InterPARES 1 Authenticity Task Force, authenticity is composed of two key elements: identity and integrity. "The *identity* of a record refers to the distinguishing character of a record, that is, the attributes of a record that uniquely characterize it and distinguish it from other records. From an archival diplomatic perspective, such attributes include: the names of the persons concurring in its formation (i.e., its author, addressee, writer, and originator); its date(s) of creation (i.e., the date it was made, received, and set aside) and its date(s) of transmission; an indication of the action or matter in which it participates; the expression of its archival bond, which links it to other records participating in the same action (e.g., a classification code or other unique identifier); as well as an indication of any attachment(s) since an attachment is considered an integral part of a record. The *integrity* of a record refers to its wholeness and soundness: a record has integrity when it is complete and uncorrupted in all its essential respects" (MacNeil et al., "Authenticity Task Force Report," op. cit., 47).

[108] In general, data lineage speaks to the provenance or origin of a particular set of scientific data, which is essential to determining the accuracy, currency and validity of derived information and any assumptions, hypotheses or further work based on those data. Expressions of data lineage are found in the metadata that accompany the datasets or in other associated documentation such as peer-reviewed papers, reports, headers or notations. Data lineage or provenance is an area of considerable discussion in the sciences. See, for example, Yogesh L. Simmhan et al. (2005), "A Survey of Data Provenance in e-Science," *Special Interest Group on Management of Data (SIGMOD) Record* 34(3): 31–36. Available at http://www.sigmod.org/sigmod/record/issues/0509/p31-special-sw-section-5.pdf.

[109] For more detailed discussion of the concepts of authenticity, accuracy and reliability in relation to the sciences in general, see the section titled "Conceptual Analysis: Authenticity, Accuracy and Reliability in the Literature of the Sciences" in the Domain 2 Task Force Report. For further discussion of these concepts in relation to the findings of the Focus 2 case and general studies, see the section titled "Authenticity, Accuracy and Reliability in the Science Focus Case and General Studies" in the Domain 2 Task Force Report.

[110] Although, as the results of the general study 09 survey demonstrates, there are certain fields in the sciences, such as archaeology, that currently are conspicuously under-serviced by data repositories. See discussions of this issue on pp. 54 and 90-91 in the General Study 09 Final Report.

articles are preserved and made available to scientists and the public by science libraries, not by science archives. Redundant paper copies of scientific publications distributed throughout the world have reduced the risk of losing the publication part of the scientific record. However, with the advent of e-journals for scientific publications and the publication of technical e-reports and conference e-proceedings on Web sites, there is an increasing risk that without e-science record archives some of these elements of the scientific record will be lost or their authenticity will be in question.

Focus 3 – Government

Research team

The group of co-investigators conducting research on the records of e-government activities included a combination of researchers drawn from the public archives environment and from the university environment. Thus, the Focus 3 Task Force was an ideal blend of practitioners and researchers. During the course of five years, representation from the national archives participating in the Project changed as the primary work commitments of individual researchers shifted. The following list includes the names of all who contributed to the work of the Focus 3 Task Force at some point over the duration of the Project.

Chair:

Philip Eppard	Jan 2002 - Dec 2006

Researchers:

Richard Blake	National Archives of the United Kingdom (withdrawn)—Working Group 3.3
Margaret Campbell	Archives of Nova Scotia (withdrawn)—Working Group 2.3
Filip Boudrez	City Archives of Antwerp, Belgium—Working Group 3.3
Hannelore Dekeyser	Katholieke Universiteit Leuven, Belgium—Working Group 2.3
Terry Eastwood	The University of British Columbia, Canada—Working Group 1.3
Fynnette Eaton	U.S. National Archives and Records Administration—Working Group 3.3
Philip Eppard	University at Albany, State University of New York, USA—Working Group 2.3
Elaine Goh	National Archives of Singapore—Working Group 3.3
Ken Hannigan	National Archives of Ireland—Working Group 1.3
Ken Hawkins	U.S. National Archives and Records Administration—Working Group 3.3
Hans Hofman	National Archives of the Netherlands—Working Group 3.3
Richard Marciano	San Diego Supercomputer Center, USA—Working Group 3.3
Terry Maxwell	University at Albany, State University of New York, USA—Working Group 1.3
John McDonough	National Archives of Ireland—Working Group 1.3
Sue McKemmish	Monash University, Australia—Working Group 3.3
Tom Quinlan	National Archives of Ireland—Working Group 1.3
Shelby Sanett	U.S. National Archives and Records Administration—Working Group 3.3
Jim Suderman	City of Toronto Archives, Canada—Working Group 3.3
Ken Thibodeau	U.S. National Archives and Records Administration—Working Group 3.3
Malcolm Todd	National Archives of the United Kingdom—Working Group 1.3

Research Assistants:

Melissa Adams	The University of British Columbia, Canada
Deidre Brocklehurst	The University of British Columbia, Canada
Natalie Catto	The University of British Columbia, Canada
Terra Dickson	The University of British Columbia, Canada
Jennifer Douglas	The University of British Columbia, Canada
Adam Farrell	The University of British Columbia, Canada
Dan Farrell	The University of British Columbia, Canada
Coby Falconer	The University of British Columbia, Canada
Fiorella Foscarini	The University of British Columbia, Canada
Joshua Hauck-Whealton	University at Albany, State University of New York, USA
Peggy Heger	The University of British Columbia, Canada
Ted Hoppenstedt	University at Albany, State University of New York, USA
Janine Johnstone	The University of British Columbia, Canada
Eleanor Kleiber	The University of British Columbia, Canada
Tracey Krause	The University of British Columbia, Canada
Andrea Lam	The University of British Columbia, Canada
Yvonne Loiselle	The University of British Columbia, Canada
Rachel McMullin	University at Albany, State University of New York, USA
Brenda McPhail	University of Toronto, Canada
Catherine Miller	The University of British Columbia, Canada
Elisheba Muturi	The University of British Columbia, Canada
Jane Morrison	The University of British Columbia, Canada
Carolyn Petrie	The University of British Columbia, Canada
Peter Runge	University at Albany, State University of New York, USA
Vincent Schillaci-Ventura	The University of British Columbia, Canada
Geneviève Shepherd	The University of British Columbia, Canada
Wendy Sokolon	The University of British Columbia, Canada
Mary Beth Sullivan	University at Albany, State University of New York, USA
Adele Torrance	The University of British Columbia, Canada
Melanie Wallace	The University of British Columbia, Canada
Carol Ward	University at Albany, State University of New York, USA
Reginald White	University at Albany, State University of New York, USA
Mark Wolfe	University at Albany, State University of New York, USA
Jessica Zacher	University at Albany, State University of New York, USA

Selection of Focus 3 case studies

Defining the precise scope of Focus 3 activity was a question that occupied researchers in the first months of the Project. Although it was clear that the general focus was to be on the dynamic, interactive and experiential records in the governmental sector, there was some discussion as to whether Focus 3 should concentrate only narrowly on activities that come under the term of e-government, which is generally restricted to the use of information technology by government to improve the delivery of services to the citizenry. Alternatively, it was suggested that the work of Focus 3 might want to cover use of such technologies in the internal workings of government as well as in the delivery of external services. In the end, the Focus 3 researchers decided to lean more toward the traditional definition of e-government.

The Focus 3 case studies were selected to represent a range of governmental levels: national, state and provincial, and city. In addition, the range of case studies was expanded with the addition of two special cases: (1) an independently developed database of international treaties and related documents and (2) a Web site maintained by an enterprise association that is similar to government information and service Web sites. This brought to eight the total number of case studies completed in Focus 3. Two additional case studies were proposed and approved, but were later withdrawn (case study 11, Nova Scotia Business Registry Service and case study 23, UK Knowledge Network). In the first instance, the case study researcher was unable to continue participating in the Project. In the second instance, the researcher ultimately could not receive the necessary permissions to carry out the study.

Summary of Focus 3 case studies

The following brief descriptions of the Focus 3 case studies identify the digital entities that were examined, while also highlighting some of the more salient findings from each study. For a more detailed contextual analysis of these case studies in relation to the records creators and the nature of the activities resulting in document creation, see the "Characterization of the Case Studies" section in the Domain 1 Task Force Report. See Appendix 3 for a complete list of participants responsible for each case study in this focus.

Case study 05: Archives of Ontario Web Exhibits[111]

This study focuses on the creation and posting of three Web exhibits within the operational contexts of two publicly funded archival institutions; the Archives of Ontario and the City of Toronto Archives. The goal of this study is to examine the processes by which Web exhibits in hopes of identifying indicators of authenticity, accuracy and reliability in experiential and interactive records. The records that the research team studied are, primarily, the final format of the exhibit records and, secondarily, the records created and used during the actual creation of the virtual exhibits. The types of records examined include scanned images, recorded sound and text files, all of which are combined into Web pages using HTML and accessed using standard Web browser applications. The basic process of creation involves decisions by Archives' staff on the subject of an exhibit, on the choice of records used to include in the exhibit, and on how to present the chosen records.

Although the Web exhibits studied here are static documents from the creator's perspective, the interaction of the rendering system (i.e., Web browser application) with the exhibits can cause considerable variance in the user's perspective of the records. Corporate standards designed specifically to diminish or minimize the vagaries of this type of interaction for Web site development exist in the jurisdictions of both institutions participating in this study. However, this study reveals that the details of the preferred or "target" rendering system are only assumed by the creator and not actually conveyed to the Web exhibit users. As such, it is suggested that, insofar as experiential digital objects are "objects whose essence goes beyond the bits that constitute the object to incorporate the behaviour of the rendering system, or at least the interaction between the object and the rendering system,"[112] the Web exhibits in this study are experiential digital objects.

[111] See Jim Suderman et al. (2004), "InterPARES 2 Project - Case Study 05 Final Report: Archives of Ontario Web Exhibits." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs05_final_report.pdf.
[112] Duranti, InterPARES 2 Project Proposal, op. cit., 1.1-3.

Some of the key indicators found to impact the authenticity, accuracy and reliability of the digital objects examined in this study include:

- the presence of a nascent or "emerging" business process that is essentially "trust-based," meaning that there is no established procedural context in terms of which individuals are to fulfil which roles in the creation process of the Web exhibits;
- a decentralized technological environment that is in a constant state of change;
- a limited or non-existent recordkeeping environment, particularly in relation to the digital records examined in this study, whose creation and maintenance was deemed to be ad hoc and at the discretion of the individuals responsible for the development and maintenance of the Web exhibits;
- legal and moral issues, particularly in relation to protecting copyright, providing accessibility to users with disabilities, upholding the requirements of donor agreements for the materials used in the exhibits, avoiding controversy in the selection of exhibit topics and source materials (censorship), and insuring the accuracy of the interpretive text;
- identification of the Web exhibits as records and not publications;
- the inability to determine the values the creating organizations place on the Web exhibit records due to the absence of any appraisals;
- spotty adherence to the InterPARES 1 benchmark requirements for supporting the presumption of the for authenticity of electronic records;
- the use of the same digital components in multiple records;[113] and
- the lack of feedback from Web exhibit users regarding their interactions with the exhibits.

Case study 12: Antarctic Treaty Searchable Database[114]

The focus of this case study is the Antarctic Treaty Searchable Database (ATSD), an online, dynamic and interactive "policy archive" of digital records, and other supporting documents, resulting from the various measures that have been approved by the Antarctic Treaty Consultative Parties from 1959 to the present.[115] Now in its 8th edition,[116] the ATSD was originally produced for educational purposes in 1999 using materials from the U.S. Department of State as part of a two-year National Science Foundation project through the National Science, Technology, Engineering, and Mathematics Education Digital Library (NSDL). The ATSD is now used in many spheres outside of education and by many organizations worldwide, including by international governmental and non-governmental organizations, national government agencies, commercial institutions and industry. The objectives of this case study include determining the functional and conceptual challenges of modeling the creation and preservation of digital records and determining the technical and administrative challenges of re-purposing public domain records.

The ATSD interfaces with a proprietary search and data integration engine, the *Digital Integration System™ (DIGIN®)* from EvREsearch Ltd.,[117] which is designed to provide "technology independent" access to the records in the system. The *DIGIN*® engine automatically extracts and integrates information from digital records (and other digital objects)—all of which are maintained in ASCII format in the system—in response to user queries. Moreover, it does

[113] For a detailed discussion of the concept of digital components and the different relationships these objects may have with records, see the Domain 3 Task Force Report. See also Duranti and Thibodeau, "The Concept of Record," op. cit.
[114] See Paul Arthur Berkman et al. (2005), "InterPARES 2 Project - Case Study 12 Final Report: Antarctic Treaty Searchable Database." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs12_final_report.pdf.
[115] See http://aspire.tierit.com/.
[116] The ATSD was in its 5th edition during the time that it was the focus of the InterPARES 2 case study.
[117] The *Digital Integration System™ (DIGIN®)* from EvREsearch Ltd. See http://evresearch.com/Integration%20Engine.htm.

this "without the use of markup or metadata that otherwise would create technology dependencies tied to "structured information" standards and legacy implementations."[118]

Although not a true recordkeeping or preservation system, the ATSD nevertheless provides a number of identifiable records management functionalities, including: records registration, classification/organization, access control and retrieval. It also appears to incorporate certain functionalities designed to facilitate long-term preservation of the records in the database.[119] Like the Atlas in case study 06, the ATSD raises some important issues regarding long-term preservation and access of authentic, multi-source digital materials maintained in interactive, Web-based systems.

Case study 17: New York State Department of Motor Vehicles On-line Services System[120]

The focus of this case study is the On-line Services System through which the New York State Department of Motor Vehicles (DMV) allows citizens to use a Web browser to conduct various online transactions related to activities such as registration renewal, driver's license ordering and renewal, custom plate ordering and driver's tests scheduling. Users conduct legal and financial transactions within the Web site, which generates records in a networked and online environment. In addition to these online transactions, the DMV's Web site allows users to find information, download forms, read publications and statistics and send e-mail to the DMV. The DMV's highly interactive online system features a complex set of interwoven electronic activities that complements the DMV's internal electronic legacy system and the business rules that have applied to that system. Only a portion of the business transactions supported by the DMV can be accomplished via its Web site, and there are no services provided by the Web site that are exclusive to that medium. Of specific interest in this study are issues relating to records creation and maintenance within the online system, issues relating to ensuring accuracy, authenticity and reliability in the DMV's recordkeeping system, and issues relating to technology.

The On-line Services System was designed in house, rather than using a commercial-off-the-shelf product; however, the DMV does use a mix of third-party and in-house products for software and security measures, including the implementation of a Public Key Infrastructure (PKI) that must conform to a strict, regulated legal framework. As a component of the New York state government, the DMV and its On-line Services System are required to comply with a number of state and federal laws, regulations and policies, including all applicable state and federal laws relating to freedom of information. Use and maintenance of the system is also subject to numerous internal DMV policies and standards.

The DMV maintains "core" records for each of its customers. A core record contains proof-of-identity information about a customer and must be created during an in-person office visit before the customer is eligible to conduct any transactions online. During an online transaction, the DMV system matches the identification information provided by the customer with the identification information in the core record file of that customer. If the two match, the information is assumed to be authentic. Because online customers only have access to transactions, not to the digital entities or the core records in the DMV's mainframe database system, and because the system also limits what type of access each DMV employee has to the digital entities in the system, the creator is "99% sure" that its records are secure from unauthorized access by users both internal and external to the system.

[118] Berkman et al., "Case Study 12 Final Report," op. cit., 41.
[119] See Ibid., 37–39 (specifically, the answer to case study research question no. 18).
[120] Philip Eppard and Mark Wolfe (2006), "InterPARES 2 Project - Case Study 17 Final Report: New York State Department of Motor Vehicles On-line Services System." Available at http://www.interpares.org/display_file.cfm?doc=ip2_c17_final_report.pdf.

Case study 18: Computerization of Alsace-Moselle's Land Registry[121]

This case study examines the creation of a computerized land registry in Alsace-Moselle, a regional administrative entity in France. The primary focus of this case study is the long-term preservation of digital signatures within a dynamic information system designed to improve the efficiency of government-citizen relations in the context of the French civil law evidence system. To this end, the study documents the efforts of the Groupement pour l'Informatisation du Livre Foncier d'Alsace-Moselle (GILFAM), the administrative body that oversees development and maintenance of the computerized registry, and the Archives de France to agree on a method whereby relevant data may be transferred to the Archives while retaining their functionality. In so doing, the study also investigates the difference between authentication and authenticity, the design and implementation of a dynamic information system in the governmental sector, and how regulatory frameworks can better harmonize with archival requirements to help facilitate the long-term preservation of digital records.

The computerized Alsace-Moselle Land Registry consists of the transcriptions and scanned images of 40,000 existing paper land registry records within Alsace-Moselle, together with the inscriptions of new real estate transactions that have occurred since the system was implemented. Various procedural and state-of-the-art security technologies and methodologies are in place to ensure the authenticity, reliability and accuracy of each entry in the database. For example, each entry requires the signature of a judge, using a PKI that combines biometric access (fingerprint scan) and digital signatures. As well, there are sophisticated logging capabilities for recording all actions and transactions that take place in the system.

Of particular interest in this system is the use of a two-staged, software-controlled procedure that enables the use of digital signatures to provide continuous authentication services, rather than the more usual one-time authentication service provided by digital signatures in most systems. Nevertheless, the GILFAM, which has a legal responsibility to provide continued access to the land registry in a fashion that preserves its evidential value regardless of technological change, has not considered the problem of maintaining the digital entities, except through the mechanisms afforded by system vendors when upgrading the database management system. Nor has it considered the problem of maintaining the evidential value of digital signatures through technological evolution.

Apart from the preservation challenges stemming from the use of digital signatures, the eventual transfer of the data in the system to an archival institution is further complicated by the fact the land registry cannot be understood in the absence of the system's dynamic and interactive capabilities for organizing the records in the system and for providing access to the information that the records contain. For this reason, it is pointless to attempt to extract and transfer stand-alone documents from the system. Moreover, because of the complexity of the system and the costs involved in attempting to duplicate it, it is not possible for an archival institution to operate its own database management system, populated with inactive records of the land registry. The study discusses two possible solutions for overcoming these preservation obstacles. The first solution involves using an XML schema to serve as a translation device between the complex data model used by the land registry so that inscriptions can be imported into a relational database sufficiently simple to be maintained by an archival institution. The second solution is for the designated archival institution to grant to the GILFAM the permission

[121] Jean-François Blanchette, François Banat-Berger and Geneviève Shepherd (2004), "InterPARES 2 Project - Case Study 18 Final Report: Computerization of Alsace-Moselle's Land Registry." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs18_final_report.pdf.

to act as the custodian of the land registry and transfer its inactive records to an archival database using the same software infrastructure.

Case study 20: Revenue On-Line Service (ROS)[122]

This case study investigates Ireland's Revenue On-Line Service (ROS), a high profile e-government service offered to tax agents and customers by the Irish Revenue Commissioners. The ROS service enables the generation, maintenance, access and preservation of electronic-based tax and other records in a secure, networked, online environment. The case study focuses on the functionality and record creating and access properties of the ROS, especially in relation to the use, via a Revenue Public Key Infrastructure (PKI), of digital certificates within ROS and how these are managed. The claims that the system meets requirements for data authenticity, accuracy and integrity are also examined, particularly in the relation to the requirements of the *Irish E-Commerce Act 2000*.

The ROS system, which has been cited as a best practice in e-government products and has received several national and international awards, operates as both a stand-alone, offline and a networked, online application that links to a Web-based Revenue portal and allows authorized customers to access relevant tax information, complete and submit tax returns and, if necessary, make or arrange payments online.

At any given time, the ROS system only maintains a subset of the main body of a customer's records. Moreover, the records in this subset are record copies; the actual authentic copies of the digital records actually reside on Revenue's Integrated Taxation Processing (ITP) back-end database system. The ITP is the central component of Revenue's overall Integrated Taxation Services framework. The ROS system's architecture separates some of the functionality of the record-making and recordkeeping systems using business rules to dictate the records contained in the subset stored within the ROS database.

Record reliability and accuracy are, in part, addressed through the use of automated validation routines built into the ROS form templates to ensure that some of the data fields populated by users are accurate and complete. Authenticity is managed through the use of digital signatures administered via a PKI. The ROS system is very tightly integrated within the Revenue PKI environment, which is used to control access to the system and to ensure the authenticity, integrity, confidentiality and non-repudiation of all transactions between Revenue and its customers. However, although the Revenue PKI facilitates the transmission of authentic data into the ROS system, the case study researchers feel that this procedure cannot, at present, be used to continue to confer authenticity. This is due to the lack of "internal controls and procedures illustrating the mechanisms by which data elements are removed from the security wrapper, ingested into the ROS system, and processed," all of which, the researchers suggest, need to be better documented, articulated and maintained.[123] The "security wrapper" is a SOAP (Simple Object Access Protocol)[124] object that encompasses the entire transaction dataset received from the customer by the ROS system. Currently, Revenue retains the security wrapper to confer authenticity and non-repudiation over time; however, there is, at present, no defined policy regarding the retention and management of the security wrappers. Moreover, it appears that the security wrappers are retained only in the ROS system and not in the central ITP system.

[122] John McDonough, Ken Hannigan and Tom Quinlan (2005), "InterPARES 2 Project - Case Study 20 Final Report: Revenue On-Line Service (ROS)." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs20_final_report.pdf.
[123] Ibid., 71.
[124] SOAP is an XML-based protocol that facilitates the exchange of information over HTTP between different applications running on different operating systems, with different technologies and programming languages.

Another key issue identified by the case study researchers is that, although creating agency maintains an “archive collection system” for certain PKI-related records such as public keys and certificates, it has not conducted a formal appraisal of the records in the ITP and its other back-end database systems, nor has it instituted a long-term archiving policy for data stored in either the ROS or the ITP back-end systems. This lack of formal appraisal and retention policies, especially with respect to the subset of ITP records and the security wrappers held in the ROS system, is identified as the most obvious deficiency in the current ROS system with respect to long-term preservation requirements.

Case study 21: Electronic Filing System (EFS) of the Supreme Court of Singapore[125]

The focus of this case study is an online filing system of records of civil matters that is meant to facilitate the process of civil litigation through e-filing and electronic information services, as well as to implement the use of digital documents in electronic chambers or electronic courts in Singapore. Through the Electronic Filing System (EFS), law firms can electronically submit and sign their legal records to the courts online at all hours; they can also serve records to other firms electronically. The EFS also supports online searching of relevant case information and the extraction of records.

Due to the complexity of the business process and juridical context that governs the creation of various types of case files created within a single system in the EFS, this case study limits its examination to the filing of bankruptcy petitions and summons in chambers bankruptcy—records that the EFS of the Supreme Court of Singapore requires be submitted in digital form—focusing on an analysis of the authenticity and reliability of the records created within this judicial e-government context. More generally, the study examines the records creation process, appraisal and preservation of the records created within EFS, and the policies, procedures and standards taken to control the creation, modification and preservation of records within the system.

The EFS is comprised of standardized HTML style sheets, XML files, Visual Basic and PDF records. Essentially, the EFS is an integrated workflow application and document imaging system. All case files received at the Court’s end are stored onto WORM (Write Once Read Many) optical disks in a jukebox.

The reliability and accuracy of the records in the EFS are addressed through the application of documentary templates and strict procedural controls that govern the creation of the records. For example, if the information entered by the law firm is inaccurate, the Court’s registry staff will reject the record and request that the law firm correct the record and re-file its submission. Authenticity is addressed through the application of various procedures, protocols and technological systems, including: an “access control matrix” that identifies the names and various roles of the action officers as well as their access rights within the EFS (the EFS does not, for example, allow records filed by law firms to be changed or modified by the Court); a PKI, together with various software and hardware encryption technologies, to secure and authenticate the transmissions of records; and various “protective countermeasures” designed to protect the integrity of the system and its records, such as firewalls, anti-virus and intrusion-detection software and periodic, intentional hacking attempts by a government agency to expose any vulnerabilities with the system.

[125] Elaine Goh (2005), “InterPARES 2 Project - Case Study 21 Final Report: The Electronic Filing System (EFS) of the Supreme Court of Singapore.” Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs21_final_report.pdf.

Case study 24: Preservation of the City of Vancouver GIS Database (VanMap)[126]

The focus of this case study is VanMap, an extensive geographic information system (GIS) containing georeferenced data on the City of Vancouver. The system is maintained by the City's Information Technology Department and incorporates data that are supplied and updated on a regular basis by the Engineering, Planning, Social Planning, Permits and Licenses, Real Estate Services and other City departments and, to a much lesser extent, by external agencies such as the provincial government and crown agencies. The goals of the case study are to understand the types of records that are generated by VanMap and to explore issues relating to security, recordkeeping and long-term preservation in relation to the use of a GIS.

Essentially, VanMap is a Web-based reference tool created to provide City staff with ready access to a detailed visual realization of the City, in the form of interactive maps, as they carry out a variety of administrative tasks. Its fundamental purpose is to meet the needs of internal City users in providing zoning, permitting, licensing, emergency planning, utilities management, traffic control, street maintenance and numerous other services to Vancouver's citizens and businesses.

The data in VanMap are constantly being updated, with the frequency of the updates varying considerably. In most cases, the updating process physically overwrites (replaces) any existing data with the new data. For data that are overwritten, there is no way to track updates over time or to access previous instantiations of the data since copies of overwritten data are not routinely kept. This process presents significant conceptual and technical challenges relating to the need both to ensure that the City government can be held accountable for the way in which the data in the system at a specific point in time are used to support decision-making activities, and to preserve the authenticity of those data and the experience of accessing them in the form of interactive maps.

The reliability and accuracy of the data in VanMap are the responsibility of the originating departments and external agencies, not the VanMap Team. There are, however, no formal data quality agreements in place between the various data providers and the VanMap Team. Instead, the Team relies on informal agreements expressed verbally or through e-mail, supported by what is characterized as an "element of trust" that exists between the Team and the various City departments. In other words, the VanMap Team trusts that the data provided by each department are reliable and accurate since, as one Team member noted, "we all work for the City."[127] Furthermore, it is noted that only a limited number of well-trained staff within the departments are able to update the data using data entry formats that are highly automated and strictly controlled. With respect to data provided by external agencies (e.g., utility agencies), these are usually accompanied by disclaimers about the accuracy and reliability of the data. In such cases, the VanMap Team is careful to post the disclaimers in association with the relevant data.

The data provided to the VanMap Team by the City departments and external agencies are not altered in any way that would affect the authenticity, reliability or accuracy of the data. Moreover, once the data are input by specified staff they are not modified by the way they are used in VanMap, since VanMap is a read-only system. Thus, the VanMap Team can, in the end, only guarantee that the data that appear in VanMap are as authentic, reliable and accurate as the original source.

---

[126] Evelyn McLellan (2005), "InterPARES 2 Project - Case Study 24 Final Report: City of Vancouver Geographic Information System (VanMap)." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs24_final_report.pdf.

[127] Ibid., 10.

Case study 25: Legacoop Bologna Web Site[128]

The focus of this case study is the Web site of Legacoop Bologna, a business association cooperative operating in the Bologna area of Italy that promotes the development of cooperation, mutual aid and solidarity, stimulates economic relationships among member cooperatives and spreads the principles and values of the cooperation ideal. The Web site provides detailed information about the cooperative, along with facts about the economic and social situation within the local territory, to its members and the general public. It also offers specialized online interactive services to its members.

Although technically a support service provided by an association of private businesses, the Legacoop Web site could fall into the category of e-government as to its function if not as to its nature. Legacoop is indeed an institutional player in the local socio-economic system and its Web site, which presents both static and dynamic information, constitutes an essential instrument for maintaining what might be called the "cooperative network" of the association's members. A restricted area of the Web site provides enhanced services to the association's members, while an unrestricted area also provides basic information and services to the general public. Further, one of the main purposes of the Web site is to increase cooperation between Legacoop's associate businesses and the city.

From a preservation viewpoint, this complex Web site presents significant challenges to preserving its authenticity over the long term. From a policy viewpoint, it provides a good starting point for conceptualizing the policy issues presented by the preservation of digital material of a private organization that behaves very much like a government. For these reasons, the case study addresses a number of interesting long-term preservation issues that are particularly relevant to the work of both the Policy Cross-domain and the Focus 3 Task Forces.

General study conducted by Focus 3

At the beginning of the Project, the Focus 3 Task Force conducted a general study in an effort to gauge how extensively governments were using the World Wide Web to provide information and services to citizens. The Government Web Sites Survey (general study 08) used a typology developed by the National Archives of Australia to classify Web sites as static, static with form-based interactivity, Web sites based on dynamic data access, and dynamically generated Web sites. A total of 321 Web sites in six countries were analyzed, with 41% judged to be static, 34% static with forms, 19% dynamic data access and 6% dynamically generated.[129] This rudimentary survey gave researchers in the government focus a sense of the terrain in which their case study research would take place. It also provided a context in which researchers could understand the extent to which interactive and dynamic environments were shaping the delivery of e-government services.

Findings specific to the governmental sector

Of the three focus areas in InterPARES 2, the governmental sector offered record creation and maintenance systems that were most similar to those studied by InterPARES 1. Nevertheless, because of the emphasis in InterPARES 2 on interactive, experiential and dynamic systems,

[128] Mariella Guercio (2004), "InterPARES 2 Project - Case Study 25 Final Report: Legacoop of Bologna Web Site." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs25_final_report.pdf.

[129] See Wolfe, "General Study 08 Final Report," op. cit.

relative to the primarily static systems examined in InterPARES 1, there were of course major differences in relation to the ways in which data and records were being generated and maintained by the creators studied in InterPARES 2. Governments carry out their core functions through the creation of records, usually under the constraints of fairly stringent guidelines or rules. The case studies carried out in Focus 3, therefore, show governments adapting their methods to these new digital environments to fulfill traditional roles with greater speed and accessibility.

These same case studies also demonstrate that there is a greater awareness among creators in the government focus of the formal elements of a document than there is among creators in the other two focuses. In fact, seven of the eight government focus case studies (the Antarctic Treaty Searchable Database being the exception) provide at least an implicit reference to the extrinsic and intrinsic elements of documentary form in relation to the records being examined, while four of the studies include explicit and often detailed discussions of the intrinsic and extrinsic elements. This stands in sharp contrast to most of the creators in the other two focuses who seem to exhibit little interest in, or awareness of, such matters. In fact, when asked to identify the key formal elements and attributes of their digital entities, creators in the arts and science focuses typically assumed the question was referring to issues such as media formats (e.g., VHS, DVDCam, etc.), file formats (e.g., .doc, .pdf, etc.) and/or the software applications that produce files in these formats, rather than to the intrinsic and extrinsic elements of documentary form that speak to the authenticity of the records.[130]

Most of the case studies in Focus 3 mainly depict governments doing traditional activities that are now carried out in digital formats. Thus, there is the question of adaptation of existing traditions to digital delivery and service mechanisms. In some cases, the new methods exist alongside older paper-based methods, acting as complements or alternatives to the older methods. On the whole, however, the purposes of creating documents have remained unchanged despite the switch or expansion to digital methods.[131]

New technologies do allow existing purposes to be augmented, however. When documents provide a service or information to citizens, the move to e-government allows for faster filing, quicker retrieval and improved access. Providing government services in an electronic environment reduces the costs of transactions while improving the speed and flexibility of the transactions. Additionally, digital security measures, such as the use of a Public Key Infrastructure,[132] allow for security and confidentiality to be safeguarded at every stage of the cycle in many cases. On the other hand, the switch to new technologies can, in some cases, create new security concerns.

As these changes represent traditional methods being carried out in a new fashion, existing methods of document creation remain largely unchanged. In most cases, only a few steps have been added to account for the new technologies used, such as digital security. In the few cases where new technologies have resulted in new methods of creation, it appears that the "newness" of the digital world obscured the fact that existing methods could still be carried out without large modifications.[133]

[130] For more detailed discussion of these findings in relation to the Focus 3 case studies, see the analysis of Domain 1 research question 3 in the Domain 1 Task Force Report.
[131] For more detailed discussion of these findings in relation to the Focus 3 case studies, see the analysis of Domain 1 research question 1 in the Domain 1 Task Force Report.
[132] See, for example, the use of a PKI and digital signatures in case studies 17 (New York State DMV), 18 (Alsace-Moselle Land Registry), 20 (Revenue On-Line Service) and 21 (Supreme Court of Singapore).
[133] For more detailed discussion of these findings in relation to the Focus 3 case studies, see the analysis of Domain 1 research question 2 in the Domain 1 Task Force Report.

Only about one half of the organizations examined in the Focus 3 case studies is actually creating records, at least according to the definition of a record as established by diplomatics and adopted by InterPARES 1.[134] In the remaining cases, some of the documents produced can be considered partial or potential records. Failure to meet record specifications frequently resulted from the use of hypermedia features, where the fluid and changing data produce a document with no fixed form that is never properly set aside. Other systems rely on providing services through dynamic databases, which are never fully fixed and set aside.[135]

An analysis of the e-government literature shows that there is awareness of and concern for the authenticity and reliability of records, although the terms are often used interchangeably. Accuracy was less of an issue, although one of the virtues of the use of digital forms for creating and entering data into systems is that a digital system can, through various automated and semi-automated controls, enforce a certain amount of accuracy that would not be easily achieved in a manual system. Authentication was a key concern in the literature and in some of the case studies, but in both cases it was sometimes taken as a guarantee of authenticity rather than a verification of authenticity at a certain point in time.[136]

The process of dealing with digital records should lead to a greater awareness of the need for records retention and preservation policies, especially because of the tendency of digital media to become inaccessible due to technological obsolescence. Ideally, this factor should encourage records creators to become more proactive in considering the retention needs of their organizations. This is, in fact, the case in some of the case studies. However, many others displayed a total lack of formalized preservation strategies. In some instances there were no long-term retention requirements on the records created—everything was considered part of the live system as long as it was needed. Even in these cases, however, the benchmark and baseline requirements set forth in InterPARES 1 should be adhered to if authenticity, accuracy and reliability of the records are to be demonstrated.[137]

General Conclusions of the Focus Task Forces

All members of the Focus Task Forces also belonged to a Domain Task Force, where their experiences with their own case study and their knowledge of the case studies being done in their specific Focus could be shared with the other researchers and inform the work being done on Records Creation and Maintenance (Domain 1), Authenticity, Reliability and Accuracy (Domain 2) and Appraisal and Preservation (Domain 3). Questions addressing the key issues in each domain were included in the basic case study questionnaire.[138] The results of the analyses of the case studies that were conducted by the domains are available in the domain reports.

---

[134] For details on how "record" is defined by traditional diplomatics and by InterPARES 1, see Duranti, "Diplomatics," op. cit.; and Duranti, *Long-term Preservation*, op. cit., especially Part 1 and Appendices 1 and 2. For a more concise summary, see the section titled "Definition of Record" in the Domain 1 Task Force Report.

[135] For more detailed discussion of these findings in relation to the Focus 3 case studies, see the analysis of Domain 1 research question 4 in the Domain 1 Task Force Report.

[136] For more detailed discussion of the concepts of authenticity, accuracy and reliability in relation to e-government in general, see the section titled "Conceptual Analysis: Authenticity, Accuracy and Reliability in the Literature of e-Government" in the Domain 2 Task Force Report. For further discussion of these concepts in relation to the findings of the Focus 3 case studies, see the section titled "Authenticity, Accuracy and Reliability in the Government Focus Case Studies" in the Domain 2 Task Force Report.

[137] For more detailed discussion of these findings in relation to the Focus 3 case studies, see the analysis of Domain 1 research question 7 in the Domain 1 Task Force Report. See also the Domain 3 Task Force Report, especially the "Findings" and "The Domain 3 Research Questions" sections.

[138] See Appendix 4.

Furthermore, each case study attempted to provide some input to the various cross-domains to supplement the research of the cross-domains. Representatives of each Focus Task Force were assigned as liaison to each of the cross-domains to ensure the free flow of relevant information among the researchers. The cross-domains were also invited to submit specific questions to be included in the case study questionnaire to learn about practices relevant to their area of study. These included questions 19 and 20 for the Policy Cross Domain[139] and questions 21 and 22 for the Description Cross-domain.[140] As a result, the Description Cross-domain received information about the existence of any in-house metadata schemas or about the adoption of existing metadata standards for use in records' control procedures by any of the case study subjects. Likewise, the Policy Cross-domain received information about legal or ethical issues underpinning current records creation practices along with pointers to other potential areas where standards were required and/or were being adopted.

Workflows were modeled for some case studies that particularly lent themselves to this form of study. The same *Integration Definition for Function Modeling (IDEF0)* methodology employed by the Modeling Cross-domain was adopted for the case studies.[141] This facilitated the comparison of actual records management practices in the field with the Project's Chain of Preservation model, an "ideal" records lifecycle model that integrates records creation, maintenance and preservation.[142] Finally, the terminology of all case studies was reviewed by the Terminology Cross-domain and selected terms were added to the dictionary section of the Project's Terminology Database to illustrate variations in the definitions of words among the many disciplines studied.

References

To explore any individual case study in more details, please consult the information available on the InterPARES 2 Web site or on the DVD that accompanies this book. In many, although not all instances, the following documents have been prepared for each case study:

- case study proposal—includes a description of the case study subject, the rationale for choosing that case study subject, the research methodologies to be used, a description of the research team and their roles, and a timeline;
- areas that should be covered—contains a standardized set of information about the case study extracted from the whole of the documentation related to the case study; used to support the characterization of the study;
- case study characterization—contains information about: (1) the creator(s) of the digital entities under study, which allowed researchers to identify who produced the digital entities and for what reasons and (2) the creator's administrative and managerial framework and about the digital entities under study, which allowed researchers to characterize the entities and the types of activities related to their creation and management;

---

[139] Question 19: Have you had to make rules, or adopt standards to help you in your work? Do you find you have to update them regularly? and Question 20: Do any legal or ethical issues arise from your electronic work?

[140] Question 21: Did you create or adopt a standard list of information which you try to records about each file, or work? and Question 22: Where did you get it? Do you know if others use the same one?

[141] For more information regarding the IDEF0 modeling methodology, see the Modeling Cross-domain Task Force Report. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_5_modeling_task_force.pdf.

[142] Ibid.

- case study overview—contains information extracted from the case study that is directly relevant to the work of all three Domains; used to facilitate analysis of case study data within each Domain;
- diplomatic analysis—contains an analysis of each type of digital object identified in the case study in relation to the five necessary characteristics of a record to determine if each object could be considered a record;
- Domain 1 analysis—contains information extracted from the case study to address the Domain 1 records creation and maintenance research questions;
- Domain 3 analysis—contains information extracted from the case study to address the Domain 1 records appraisal and preservation research questions; and
- final report.[143]

To locate conference presentations and articles addressing specific case studies, or Focus-specific analysis of case studies, please consult the list of participants, listed by case study, in Appendix 3. The contributions of the various investigators can be located in the extensive database where the Dissemination Activities of Project members are recorded. Again, this is available both on the InterPARES 2 Web site[144] and on the DVD that accompanies this book.

For example, the principal investigator for case study 12 (Antarctic Treaty Searchable Database) is Paul Berkman. A search of the Dissemination Activities database using the keyword "Berkman" will locate citations for two conference papers written by Berkman and citations for two refereed articles co-authored by Berkman, one of which is:

> Berkman, Paul Arthur, George James Morgan III, Reagan Moore and Babak Hamidzadeh. 2006. "Automated Granularity to Integrate Digital Records: The "Antarctic Treaty Searchable Database" Case Study," *Data Science Journal* 5: 84–89. http://www.jstage.jst.go.jp/article/dsj/5/0/84/_pdf

Bibliographies

The bibliographies developed for each case study are available in their respective final reports. In addition, these bibliographies were merged with more extensive bibliographies (including, in some cases, annotated references) developed by the domains and cross-domains in their research. This integrated bibliography is available on the InterPARES 2 Web site[145] and on the DVD that accompanies this book.

[143] See Appendix 23 for a comprehensive listing of all major case and general studies-related documents organized by resource category.
[144] See http://www.interpares.org/ip2/ip2_dissemination.cfm?proj=ip2.
[145] See http://www.interpares.org/ip2/ip2_documents.cfm?cat=biblio.

PART TWO

RECORDS CREATION AND MAINTENANCE

Domain 1 Task Force Report

by

Martine Cardin, Université Laval

Introduction[1]

Background and mandate

An understanding of the meaning of a record rests upon an understanding of the process of its creation and of the function of the record within the activity in which it participates. To preserve such a record in authentic form over time, it is necessary not only to know its characteristics, processes of creation and function, the purposes for which it is kept by its creator and which of its intrinsic and extrinsic elements can be used to evaluate its authenticity, but also to ensure that the record is generated in such a way that it is possible to carry it forward for use by future generations. This knowledge needs to be acquired and developed across the wide spectrum of digital records identified in each focus area of this Project. Domain 1 was founded on this perspective. Its objective was to explore records and the processes that create and maintain them.

Although the creation process and documentary form of records created by governments tend to be regulated and controlled, thus making it easier for a preserver to carry such records forward, the electronic delivery of government services, which is redefining the processes by which transactions between a government and its citizens are conducted, may be changing not only the form of the resulting records, but also their other salient characteristics (for example, their fixity). The very nature of records created in the course of an online interaction may be very different from that of digital records presently generated in databases and document management systems, and their function may significantly differ from that of the records examined in InterPARES 1.

Records generated outside of government are largely the product of unregulated processes and have already posed interesting challenges to those responsible for their preservation. Visual artists, musicians and choreographers accumulate material with great cultural value (for example, sketches and drafts) that corresponds to the traditional definition of a record, as do organizations and individual scholars who carry out scientific research. When these materials are on paper, they pose few problems for preservation, because they are kept in their original immutable form, which remains equally accessible through time and, in most cases, facilitates determination of their identity and integrity, regardless of labelling conventions, archival descriptions, etc. In the last two decades, however, digital environments have changed the practices of artists and scientists. In some cases, these environments provide comprehensive traces of creative processes that used to go undocumented, so that we now have records of activities that were never recorded before. In other cases, the opposite has occurred. Few of the documents produced using new information technologies have properties that allow users to determine their identity and integrity, to arrange and describe them, to ensure that they can be kept accessible, and to ensure that their authenticity can be maintained and subsequently verified through time. It is important to understand how digital work environments have changed the process of record creation in each of the activities in question, how the identity of the various types of records created can be established as to provenance, authorship, function and relation to the records participating in the same activity, if and when records created in such environments can be considered complete and/or capable of accomplishing the purposes for which they were generated, and what are the criteria and practices of their creator in maintaining them.

[1] This report was carried out under the direction of Professor Martine Cardin with the assistance of Peter Gagné, student in Archival Studies at Université Laval.

Research team

The following is a list of researchers and research assistants who participated in the Domain 1 Task Force throughout the Project:[2]

Chairs:

Malcolm Todd	Feb 2004 - Sept 2004
Martine Cardin	Sept 2004 - Dec 2006

Researchers:

Paul Berkman	University of California, Santa Barbara, USA—Working Group 1.2
Martine Cardin	Université Laval, Canada—Working Group 1.1
Henry Daniel	Simon Fraser University, Canada—Working Group 1.1
Luciana Duranti	The University of British Columbia, Canada—All Working Groups
Barbara Craig	University of Toronto, Canada—Working Group 1.2
Henry Daniel	Simon Fraser University, Canada—Working Group 1.1
Terry Eastwood	The University of British Columbia, Canada—Working Group 1.3
Ken Hannigan	National Archives of Ireland -Working Group 1.3
Michael Longton	University of Victoria, Canada—Working Group 1.1
Terrence Maxwell	State University of New York at Albany, USA—Working Group 1.3
Michael Murphy	Ryerson University, Canada—Working Group 1.1
Andrew Rodger	Library and Archives Canada—Working Group 1.1
Fraser Taylor	Carleton University, Canada—Working Group 1.2
Malcolm Todd	The National Archives of the United Kingdom—Working Group 1.3

Research Assistants:

Natalie Catto	The University of British Columbia, Canada
Seth Dalby	The University of British Columbia, Canada
Heather Dean	The University of British Columbia, Canada
Heather Daly	The University of British Columbia, Canada
Jennifer Douglas	The University of British Columbia, Canada
Ann Forman	The University of British Columbia, Canada
Peter Gagné	Université Laval, Canada
Jessica Glidewell	The University of British Columbia, Canada
Nadine Hafner	The University of British Columbia, Canada
Keum Hee Yu	The University of British Columbia, Canada
Tracey Krause	The University of British Columbia, Canada
Tracey Lauriault	Carleton University, Canada
Philippe Perron	Université Laval, Canada
Carolyn Petrie	The University of British Columbia, Canada
Julie Simard	Université Laval, Canada
Geneviève Shepherd	The University of British Columbia, Canada
Frédéric Smith	Université Laval, Canada
Sherry Xie	The University of British Columbia, Canada

[2] Researcher membership in Domain 1 changed somewhat over the five years of the Project. Among those who were interested in Domain 1 issues but were unable to participate for the full length of the Project are: Paolo Buonora, Archivo di Stato, Italy; Su-Shing Chen, University of Florida, USA; and Susan Kennard, Banff New Media Institute, Canada.

Research Methodology

The three working groups within the Domain 1 Task Force were charged with examining the central concepts relating to records and record creation and maintenance processes. To achieve this, the investigators used a multi-method research approach involving three main analytical exercises: 1) analysis of case studies, 2) diplomatic analysis and 3) modeling.

First, using grounded theory, Domain 1 analyzed the case studies conducted across a wide spectrum of activities to gather information about record-making and recordkeeping processes and the records resulting from them. To achieve this purpose, it developed several tools to gather data from the case studies and produced overviews relevant to the work of all three domains. Such overviews served to highlight and summarize the Domain 1 issues as they related specifically to the findings of each case study and applicable general study, which, in turn, helped researchers to understand and evaluate the general guidelines that have emerged as a partial result of these findings.

Second, Domain 1 used the diplomatic analysis done on each case study to describe the formal elements of the records and their processes of creation and to identify the pertinent contextual information that needs to be preserved. In InterPARES 1, diplomatic analysis was used to identify the records among all types of recorded information present in each case study and to ascertain the extent to which traditional record elements continue to appear in digital records, by comparing unknown realities against the known one; that is, against the ideal template of the traditional record. In InterPARES 2, the approach was that of the original diplomatists: an examination of a wide variety of records served to identify elements, attributes and their functions and to generate templates reflecting the abstract forms of experiential, dynamic and interactive records by identifying the necessary characteristics of each of those records: that is, all the possible elements and attributes distinguishing each.

Finally, the original plan was to represent the records observed during the case studies in models and to test them against the templates to see whether the key characteristics necessary for the authentic preservation of the records are consistent across activities. In addition, the plan was to abstractly represent the creation and maintenance processes of each type of record reflected by each template using activity modeling and then generate workflows that could be tested against the typical processes of each activity. However, at the first InterPARES 2 plenary workshop, it became clear that the modeling would not be limited to the documents' creation and maintenance contexts. Instead, the modeling would cover the entire range of activities and therefore also the entire range of issues covered by the other domains of the Project. It is for this reason that the International Team decided to create the Modeling Cross-domain Task Force, which would be specifically dedicated to modeling.

Activities carried out by the Task Force over time

The meetings of the Domain 1 Task Force started in February 2004. The delay in beginning these working meetings is explained by the fact that during the first two years of the Project, research was essentially carried out by the researchers in the three focus task force groups through the production of case studies. In September 2004, the first group of nine case studies was completed, with several others still underway. Domain 1 undertook an analysis of the accumulated material. The first order of business was to come up with a way to document the work produced to be able to compare the types of activities and creative entities being investigated. It was in response to this need that a first research tool was produced to support the characterization of each case study: a template identifying a set of information common to each case study but spread over the narrative of the final reports.[3]

The template was divided into two sections. The first concerned the Creator of the Records and the second the Administrative and Managerial Framework. The information the Domain 1 Task Force was looking for about the creator was largely embedded within the contexts section of the final case study reports (i.e., section "C" in the InterPARES 2 Reporting Framework).[4] Details about the provenancial, juridical-administrative, procedural, documentary and technological contexts allowed the Task Force to understand who produced the digital records and for what reasons. The second section of the template was split into two sub-sections: Administrative and Managerial Framework and Digital Entity/Entities Under Study. Both sub-sections aimed to gather information to allow the characterization of the types of activities and entities that were being studied.

After being presented and discussed by the International Team at the Syracusa meeting in December 2004, the template was applied to each case study by a team of research assistants, resulting in an "Areas that Should Be Covered" report for each case study.[5] The necessary information could come from the final or interim case study report or from other available documentation, such as the creators' Web sites. Other information was inferred, based on researchers' knowledge of certain fields (for example, the management structure of the Archives of Ontario). A brief description of each item in the template was accompanied by references to the source material. These descriptive reports were the basis for a narrative characterization that was generated for each case study.[6] Once completed, the template, together with the accompanying instructions and the text of the characterization, was sent to the appropriate case study team leader for validation. If any item was unavailable or inapplicable to the case study, the team leaders were invited to explain why this was so, wherever possible.

At the end of this process, the Domain 1 Task Force had a tool that assisted the research in several ways. First, it permitted the validation of the case studies that had been completed up to that point and highlighted certain points that needed to be further developed in the impending final report. In short, it helped to fill in certain gaps in the reports and was useful as a methodological guideline for writing the final reports of the cases studies that were still underway. It should be noted that this was not an additional questionnaire, but a sort of reminder list of details to be included in various aspects. Secondly, this tool served as a sort of practical index for the researchers in the different groups, who could use the references to return to the source material as needed to clear up a given question. Lastly, by characterizing all of the case studies on a common basis, it was possible to make solid comparisons between them.

This characterization exercise was carried out continuously throughout the Project, as each final report became available. The production of the document, "Case Studies at-a-Glance,"[7] a sort of synthesis of the work, allowed researchers to follow the progress of the work. This document allowed the characterizations to be associated with the analyses done by other groups, including the diplomatic analyses[8] done by a team from the University of British Columbia and

[3] See Appendix 10.
[4] See Appendix 8.
[5] The "Areas that Should Be Covered" report for each case study is available on the InterPARES Web site (http://www.interpares.org/) and is also included on the DVD that accompanies this book.
[6] The "Characterization" report for each case study is available on the InterPARES Web site and is also included on the DVD that accompanies this book.
[7] See Appendix 11.
[8] The "Diplomatic Analysis" report for each case study is available on the InterPARES Web site and is also included on the DVD that accompanies this book. Note: no diplomatic analysis report exists for case study 08 (Mars Global Surveyor Data Records in

the walkthroughs of earlier versions of the Chain of Preservation Model done by the Modeling Cross-domain.[9] Finally, an annotated bibliography specific to creation and maintenance was produced.[10]

A comparative data analysis was then conducted, based on this body of accumulated material. Its aim was to determine the limits of the study sampling and to identify the presence of patterns in creation and maintenance. As a result, a preliminary overview, which aimed to analyze and characterize the entire body of case studies completed by researchers, was presented at the plenary workshop in Vancouver in September 2005. The relevance of such an exercise was to:

- *Critique the sources*. Because the case studies were important sources (although not the only ones) on which InterPARES 2 products would be based, it was essential, from a methodological perspective, to have a critical view of them. Thus, the presentation was a way to get input and validation from the researchers. It also allowed the Task Force to refine the analysis parameters.
- *Provide a common basis for the guidelines*. The case studies have observed how digital records are created, maintained and used in various milieus. On that basis, from an ethnographic perspective, the overviews profile certain particular schemas of cultural practices related to the production of digital entities. Such knowledge could scientifically support the conception of guidelines and their dissemination to various communities.
- *Support the dissemination of the case studies*. The case studies are more than a collection of data for InterPARES 2 researchers. In fact, they are considered as distinct InterPARES 2 research products themselves and could be used as tools for educational purposes. The only caveat is that, in spite of the fact that reports shared a standard reporting framework and that each has been summarized in a standard way, they are still raw material. Thus, they needed to be linked in a consistent narrative.

The Domain 1 Task Force undertook a second phase of analysis in the fall of 2005. Based on the defined parameters, the researchers reviewed the set of questions specific to Domain 1[11] and produced a set of instructions for extracting answers from the final case study reports. This extraction was carried out by a team of UBC research assistants under the supervision of Bonnie Mak and Terry Eastwood. The responses extracted from each case study were than compiled and analyzed by a team from Université Laval under the direction of Martine Cardin, culminating in a document that synthesized the information on the practices of records creation, recordkeeping, metadata schema and technology used by records creators in each of the case studies.[12]

---

the Planetary Data System); a draft analysis was conducted by the author of the case study report, but the draft analysis was never validated by the researchers responsible for the diplomatics analyses.

[9] See William Underwood, Kevin Glick and Mark Wolfe (2007), "InterPARES 2 Project - General Study 12 Final Report: Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs12_final_report.pdf; and Randy Preston (2004), "InterPARES 2 Project - Modeling Cross-domain: Walkthrough of the Manage Chain of Preservation Model Using Case Study 14 Data," draft report. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs14_COP_model_walkthrough.pdf.

[10] This bibliography was later merged with similar bibliographies developed for each case study, the other two domains and each of the cross-domains. The integrated bibliography is available on the InterPARES 2 Web site at http://www.interpares.org/ip2/ip2_documents.cfm?cat=biblio and on the DVD that accompanies this book.

[11] See Appendix 9.

[12] The "Domain 1 Research Questions" analysis report for each case study is available on the InterPARES Web site and is also included on the DVD that accompanies this book. Note: no Domain 1 analysis report exists for case study 17; however, answers to the Domain 1 research questions are provided on pp. 22–23 in the case study's final report.

Once this work was completed, a series of overviews was written for each case study.[13] The aim was to get relevant and significant quotations from the case study final reports and other validated material about specific aspects of creation and maintenance. Because maintenance and preservation issues are sometimes related, the Task Force looked at them together. For the same reason, despite the fact that they are Domain 2 issues, the Domain 1 Task Force also considered authenticity, reliability and accuracy aspects. However, in both cases the Task Force tried to keep the focus on Domain 1 issues. In addition, overviews of four InterPARES 2 general studies were generated for the same purposes. This stems from a decision at the Singapore International Team meeting in June 2006 to examine the general studies and pull out information relevant to Domain 1 for the Domain 1 Task Force Report. Consequently, Domain 1 examined the following four general studies: (1) general study 03, Preserving Interactive Digital Music: The MUSTICA Initiative; (2) general study 04, Survey of Recordkeeping Practices of Composers; (3) general study 07, Survey of the Recordkeeping Practices of Photographers Working with Digital Materials; and (4) general study 08, Survey of Government Web Site Interactivity. Such overviews served to highlight and summarize the Domain 1 issues as they relate specifically to the findings of each case study and applicable general study, which, in turn, will helped researchers better understand and evaluate the general guidelines that emerged partly as a result of these findings.[14]

Finally, a first draft of the Domain 1 Task Force Report was written and presented for discussion at the Los Angeles plenary workshop in September 2006. The draft report was then modified and augmented based on the comments received, as well as on further review of the InterPARES 2 documentation and as newer versions of some documentation became available.

Outcomes

The main outcomes of the research conducted in Domain 1 are: (1) the Template for Case Studies Analysis, (2) the set of case study characterizations, (3) the set of case study overviews and (4) the general overview on creation and maintenance observed in the larger context.

All of these research tools and products have been posted on the InterPARES Web site and communicated in presentations, lectures and scholarly writings. In addition, the Domain 1 work of collecting and analyzing data, and of developing methodology statements, coalesced with the research conducted in Domain 2 and resulted in the development of guidelines that can be used by various kinds of records creators to produce and maintain records that can be authentically preserved over the long term.[15]

Conceptual Basis

This section presents the theoretical and conceptual foundations of Domain 1. Before delving into the concepts particular to this domain, it first presents an overview of the fundamental concept of “record” as it evolved throughout the first and second phases of the InterPARES Project. This introductory presentation is a distillation of the article by Luciana Duranti and Ken

[13] The “Overview” report for each case study is available on the InterPARES Web site and is also included on the DVD that accompanies this book.
[14] The “Overview” report for each of the four general studies examined by Domain 1 is available on the InterPARES 2 Web site and on the DVD that accompanies this book.
[15] See *Creator Guidelines* in Appendix 20.

Thibodeau on the subject.[16] Following this foundation, the two notions essential to Domain 1 are presented; that is to say, the notions of creation and maintenance. These terms are defined as they are understood and used in the context of InterPARES 2, and, in particular, in the varied documentation produced by Domain 1. Finally, to situate these definitions in their archival context, the notions of creation and maintenance are presented as they apply to the concept of the records lifecycle.

Definition of record

The concept of "record" is based on that of a document as understood in archival science. In the simplest of archival terms, a document is recorded information. In turn, a record is a document that is created (made or received) as an instrument or by-product of a given activity and that is set aside for action or reference. Therefore, what distinguishes a record from a document that is not a record is the fact that a record has a relationship with the activity of the creator and is produced in the course of carrying out the activity. Research in the first phase of the InterPARES Project[17] sought to determine the soundness of the above definition for records mandated for accountability and administrative needs created in databases and document management systems.

InterPARES 1 Project

The definition of a record used by the InterPARES 1 Project is that which has been adopted by traditional diplomatic analysis and is based on a consideration of the dual notions of the identity and integrity of a document. In this definition, consideration is given both to the characteristics of a document itself (fixed form, stable content, etc.) and to the five contexts of its creation (juridical-administrative, provenancial, procedural, documentary and technological), to generate a "statement" of the presumption of the document's authenticity.

InterPARES 1 further defined the term "electronic record" as a record that is set aside and used in electronic form, regardless of the original form in which it may have been made or received. For example, a scan of a record originally in paper form is considered an electronic record. InterPARES 1 researches wanted to include digital records in the fundamental assumption of diplomatics that, regardless of differences in nature, provenance or date, from a formal point of view, all records are similar enough for it to be possible to conceive of a single document template containing all possible elements of a record. By extension of this principle, the InterPARES 1 researchers posited that the same formal elements found in traditional records could be found in digital records and that all digital records share the same formal elements. The researchers thus created a template of the four elements of an electronic record: documentary form, annotations, context and medium. The last element, medium, should not be considered one of the record's necessary constituent parts, however, but rather part of the record's technological context.

The definition of "record" by InterPARES 1 differs significantly from the traditional diplomatic definition, due to a fundamental difference in the way InterPARES 1 defines "document," which, as noted above, forms the basis of the concept of "record." While InterPARES 1 defines a document simply as "recorded information," diplomatics defines a

[16] Luciana Duranti and Kenneth Thibodeau (2006), "The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES," *Archival Science* 6(1): 13–68. (Note: a reprint of this article is included in Appendix 2.)
[17] See http://www.interpares.org for more information on the InterPARES Project and its two phases.

document as "information affixed to a medium." This difference is of great importance since "in the digital environment, it is possible to generate something that to all appearances is a document, but is not affixed to a medium."[18] This situation is based on the difference between the form in which a document is stored digitally (i.e., its binary form) and the form in which it is manifested to a person (i.e., reconstituted from its digital components and presented in a human-readable format on a computer screen or other display device), which distinguishes a digital document from a traditional one. It also leads to the distinction of the terms "stored digital record," which is a digitally-encoded object that is managed as a record,[19] and "manifested digital record," which, effectively, is defined as a digital document, treated as a record, that is visualized or rendered from a stored digital record and/or stored digital component(s) in a form suitable for presentation either to a person (i.e., in human-readable form) or to a computer system (i.e., in machine language).[20]

InterPARES 2 Project

The InterPARES 1 definition of a record worked well for the digital environments studied within the bounds of the first phase of the Project. However, when the Project expanded in its second phase to include documents created as by-products of artistic, scientific and governmental activities in interactive, experiential and dynamic environments, the question arose whether the traditional concept of a record needed to be refined or revised due to the unique characteristics of the environments being studied. In other words, InterPARES 2 research sought to determine whether the digital entities created and maintained in these environments indeed are, or could be, records based on the definition retained by InterPARES 1, or whether these entities have unique characteristics, due to the interactive, experiential and dynamic environments in which they are created, that would force a re-consideration or revision of the InterPARES 1 definition of "record."

Consequently, InterPARES researchers, Luciana Duranti and Kenneth Thibodeau, examined the aptness of the concept of record adopted by InterPARES 1 as it pertains to the InterPARES 2 case studies, particularly with the aim of examining "whether there are differences in the nature of the records produced in environments that only exist in the digital domain."[21] In some of these environments, the system itself acts as an agent for the system owner, carrying out individual transactions and creating a digital record without any physical, real-time input from the system owner or user.

Duranti and Thibodeau concluded that the concept of record did indeed need to be revised to include those records created and maintained in interactive, experiential and dynamic environments, particularly in environments that make use of digital technologies in innovative ways. Despite the fact that the electronic environments studied in InterPARES 2 can produce the digital equivalents of traditional documents, when the existing electronic document template was applied to the InterPARES 2 case studies, only slightly more than half of the systems studied were deemed to contain records. What is more, when the given systems did contain records, what records they did contain rarely resembled the model represented by the InterPARES 1 template.

[18] Duranti and Thibodeau, "The Concept of Record," op. cit., 27.

[19] The literal definition of stored digital record is given in the InterPARES 2 Terminology Database as "A stored digital document that is treated and managed as a record." However, when taking into account the embedded concept of a digital object, which comprises a stored digital record, the effective definition becomes "A digital object, placed in a storage system on a digital medium, that is managed as a record, and which includes information about the properties of the object and may also include methods of performing operations on or with the object" (http://www.interpares.org/ip2/ip2_terminology_db.cfm).

[20] Ibid.

[21] Duranti and Thibodeau, "The Concept of Record," op. cit., 22.

That was because, in addition to digital documents that mirror the form of traditional documents, these systems can also create documents that either differ significantly from traditional documents or have no traditional counterparts.

After examining the documents created in interactive, experiential and dynamic systems, Duranti and Thibodeau came up with a taxonomy of static and interactive documents based on the classes of documents created in the various types of systems. Their results are presented in an abbreviated form below.

Table 1. Taxonomy of Static, Interactive and Dynamic Documents[22]

Class	Description
1	**Static Documents**: Digital documents are static when they do not provide possibilities for changing their manifest content or form beyond opening, closing and navigating within the document.
1.1	The electronic equivalents or counterparts of traditional documents.
1.2	Documents that have no exact counterpart in hard copy or analogue form, but have fixed documentary form and content.
2	**Interactive Documents**: Documents that present variable content, form or both whose rules governing the context and form of presentation may be either fixed or variable.
2.1	Non-dynamic Interactive Documents: Documents where the rules that govern the content and form of presentation do not vary, and where the content presented in any instance is selected from a fixed store of data within the system.
2.2	Dynamic Interactive Documents: Documents where the rules that govern the content and form of presentation *may* vary.
2.2.1	Documents where the content and/or its presentation vary because it includes or is otherwise impacted by data that change frequently.
2.2.2	Documents where the content varies because it includes data received from external sources and not stored within the system.
2.2.3	Documents created in dynamic computing applications, which select different sets of rules—software applets or service components—to create the documents depending on variations in user inputs, in the sources of content data, and in the characteristics of that content.
2.2.4	Documents created by adaptive or evolutionary computing applications, where the software that generates the documents can change autonomously.

As can be seen from the above table, digital documents are divided into two main categories: static and interactive. It can also be seen that dynamic documents (Class 2.2) are a subset of interactive documents. However, which of the classes of documents can be considered records?

Fixed form and stable content are two of the key characteristics of an electronic record, as defined by InterPARES 1. Documents that fit into Class 1 presented in the above table therefore have the fixity of form and stability of content necessary to be considered records, provided that the other criteria are met. For the other classes of documents, the content or form may vary, but, for some, the type of variance may not necessarily prevent the documents from being considered

[22] Adapted from Table I in Duranti and Thibodeau, "The Concept of Record," ibid., 45–46.

records. For example, "there may be variations in the manifest form and/or content, even when there is no variation in the stored digital data used to generate the manifested document."[23] As such, some variations in form or content do not automatically prevent certain classes of interactive or dynamic documents from being considered records.

Certain documentary forms include variable elements that are controlled or intended by the author and that alter the form or presentation by allowing variable subsets of the content to be displayed at any moment. Although such cases present the user with what appears to be a document, this display is actually only a subset or part of the existing document, such as an online catalogue. However, "cases where the documentary form permits selective display of subsets of the content can satisfy the requirements for fixed content."[24]

If an interactive or dynamic document's form or content varies according to variable rules, this variability will prevent the documents from being considered records precisely because their content and/or form is not fixed. Nonetheless, it may be possible for digital documents in which fixed rules govern variations in content and/or form to be considered records, or as drafts in the process of being developed (i.e., potential records).

In their examination of the various types of records in interactive, experiential and dynamic systems, the authors also identified digital documents that, based on their use, can be described as "enabling documents."[25] Found primarily in the case studies in the arts focus of InterPARES 2, although present in other domains, these documents enable the presentation of interactive visual art or music, for example. There are at least two sub-types of enabling documents: 1) instructions for executing or producing a performance and 2) descriptions of the components, context, preconditions or requirements for performance, whose execution allows for future performances. The first sub-type of enabling documents may, for example, be likened to the script of a play, the scenario of a film or a musical score, and describe how all of the components fit together to execute the performance, while the second sub-type is like a detailed description of all the actors, props, locations, etc., used in the performance. InterPARES 2 arrived at the conclusion that, in the arts focus at least, these two types of enabling documents—the set of instructions and related information to carry out the instructions—are the necessary means for reproducing or re-creating digital artwork and music, and should be distinguished from the documents of the performance itself.

With the exception of documents where changes to content data result from system changes or the failure to retain data in the system—not through the explicit intention of the author—all interactive documents are enabling documents. Certain of these enabling documents may qualify as records. This occurs in certain situations where the document gathers some or all of its data from external sources. The requirement of having a fixed message is not met if the external data are not stored concurrently with the digital components of the record. However, "a document that delineates a fixed form in which external data are to be presented and may include some unaltered content may be an instrumental or instructive record […] The record in such cases is the digital entity, not the human-perceivable form which is reproduced from it."[26]

This last statement is quite different from the definition of an electronic record used in InterPARES 1, in which an electronic record is one that was manifested by a computer system to a human or another system. In other words, for InterPARES 1, the electronic record was the

[23] Ibid., 24.
[24] Ibid., 26.
[25] Ibid., 34–35.
[26] Ibid.

manifested record, not its stored digital component(s). The stored digital component(s) were seen to enable reproduction of the record, but were not considered the record itself. Duranti and Thibodeau's research has lead to a different, more nuanced and inclusive concept of a digital record and "to the recognition that a digitally stored record includes not only the data which must be processed in order to reproduce the manifest record, but also the rules for processing the data, including rules which enable variations in the content or form of the manifest record."[27] In this view, the digital components themselves may be seen as a record or a set of records, depending on how the different types of data (content data, composition data, form data) are instantiated in the system. "What is essential is that the computer stores and processes the data and the instructions in a way that consistently and correctly distinguishes each type and combines the different digital components of a record."[28] In other words, for the manifested document to be considered a record, it must be possible to reproduce it repeatedly as it appeared the first time.

Duranti and Thibodeau also make the distinction between retrospective and prospective records. Retrospective records fulfil the traditional, memorial function of records to bear witness to or remember a past action in which they participated or of which they were the by-products. Prospective records add a new and different dimension to the role of records. Rather than witnessing the past, they guide the future though a set of instructions or actions *to be carried out in the future*. In other words, they are the enabling records that were discussed saw above. "Retrospective records capture, while prospective records enable or at least inform interactions, experiences or dynamic processes."[29] These new types of records that have emerged in the interactive, experiential and dynamic systems in the case studies of InterPARES 2 bring an entirely new dimension to the concept of record.

InterPARES 1 has already affirmed that long-term preservation of digital records is not possible, only the ability to reproduce these records accurately, authentically and reliably. What prospective records add to the realm of digital records is the ability to reproduce them in the future, due to the set of instructions that they contain and the description of the component parts of the performance or action for which the records were originally created. The concept of "record" as now defined by InterPARES 2 is fully in keeping with the previous concept that was adopted; it merely expands upon this definition and brings digital records into the future.

Record creation and maintenance

Having defined the concept of "record" according to InterPARES 2, the discussion now turns to how the Project studied these records and the role that Domain 1 played in this research.

The overall aim of the second phase of the InterPARES Project was "to develop and articulate the concepts, principles, criteria and methods that can ensure the creation and maintenance of accurate and reliable records and the long-term preservation of authentic records in the context of artistic, scientific and governmental activities that are conducted using experiential, interactive and dynamic computer technology."[30] This mandate is broken down into three domains,[31] alluded to in the above quotation: Creation and Maintenance; Authenticity,

[27] Ibid., 27.
[28] Ibid., 30.
[29] Ibid., 33.
[30] InterPARES 2 Project Summary. Available at http://www.interpares.org/ip2/ip2_index.cfm.
[31] See InterPARES 2 Intellectual Organization. Available at http://www.interpares.org/ip2/ip2_intellectual_organization.cfm. There are also four cross-domains that address questions pertinent to all areas of inquiry in the project: Terminology, Policy, Description and Modeling.

Accuracy and Reliability; and Methods of Appraisal and Preservation. Each of these three domains cross the three focuses or "contexts" as described above, which are the artistic, scientific and governmental activities of the case study subjects.

Needless to say, the domains are not isolated or exclusive. Creation and maintenance includes considerations of authenticity, accuracy and reliability,[32] which may also be found in appraisal and preservation issues. Each domain is connected to and informs the two other domains and is in turn informed *by* them.[33] In addition, the Policy and Description Cross-domains, by their nature, obviously interrelate the three domains of InterPARES 2, as exemplified by the Description Cross-domain, which sought, in part, to determine the role of descriptive schemas and instruments in records *creation*, control, *maintenance*, appraisal, preservation and use in emerging recordkeeping systems in digital and Web-based environments in the three focus areas.[34]

Domain 1, Creation and Maintenance, studied the nature of records and of the processes that create and maintain them. As seen in the Methodology section of this report, Domain 1 made use of a series of seven questions[35] that focus on issues such as the purposes of document creation, the processes that result in document creation, document elements and attributes, the applicability of the current definition of "record" to the documents of each case study, the capture of documentary evidence, the responsibilities and liabilities related to the use of these documents and the determination and implementation of record retention decisions. These questions touch upon various elements of "creation" and "maintenance" as they are understood in and apply to InterPARES 2.

Creation

"Creation" is a term that may seem obvious at first glance. To create is to make something; give it form or life. In the case of traditional paper-based records, "creation" often meant writing, typing or otherwise physically applying information to a support—the *process* by which a record was made. In the electronic environment—especially in interactive, experiential and dynamic systems—the processes by which records are created often resemble the traditional record creation process very little or not at all. In many cases, "The use of digital technology to create records has…allowed for the bypassing of procedural controls"[36] that had been put in place to guide and define records creation. How, then, can "creation" be defined in such a context?

The InterPARES 2 Terminology Database defines "created record" as follows: "A made or received document declared a record and set aside for action or reference." Two important points can be gleaned from this definition. First, in the archival sense, creation does not simply imply making a record, but also the act of setting it aside. This organic and automatic setting aside of records is what distinguishes them as archival documents. Secondly, creation not only applies to the *making* of records but also to the *reception* of records,[37] along with setting them aside, as has

---

[32] See Records Creator Principle C4 in Appendix 19.

[33] For an example of the interrelation of creation and preservation as pertaining to digital records, see Records Creator Principle C7, ibid.

[34] Anne Gilliland (2005), "Discussion Paper on the Nature and Role of Metadata in the Creation of Reliable and the Preservation of Authentic Records in Electronic Systems," paper presented at InterPARES 2 Project Plenary Workshop 13, 20-24 February 2005, Vancouver, BC, Canada, 2 (unpublished). Emphasis added.

[35] A consolidated list of the Domain 1 research questions is provided in Appendix 9.

[36] Luciana Duranti (2001), "International Research on Permanent Authentic Records in Electronic Systems (InterPARES): Experiential, Interactive and Dynamic Records," SSHRC MCRI InterPARES 2 Project Proposal, 412-2001, 1. Available at http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf.

[37] Technically speaking, what is received is a document. It becomes a record only after being "declared" as such by being registered and classified by the recipient—a process that constitutes the *intellectual* setting aside of the record—followed by the

already been discussed. This is an important point, for it allows documents made by a third party to be included in the fonds of a given creator.

The term "records creator," intrinsically linked to that of "creation," is also found in the InterPARES 2 Terminology Database, which defines it as "The physical or juridical person who makes, receives or accumulates records by reason of its mandate/mission, functions or activities." Included here, once again, are the twin notions of making or receiving records. "Accumulating" may appear to be a third notion involved in creation, but it is merely an expression of the notion of "setting aside," as we have seen in the definition of "creation." The reference to the mandate, mission, functions and activities of the creator is another expression of the organic nature of this accumulation.

Specific creators exemplified in the InterPARES 2 Project are identified among the various stakeholders that the Project identified, and include:

- *Individual records creators*, who rely on records for continuing use, reference purposes, cultural purposes, to carry out other activities, as evidence of their work or as proof of individual rights;
- *Organizations*, which rely on accurate, reliable, and authentic records to carry out their business, fulfil legal obligations, understand previous activities and ensure continuity; and
- *Governments*, which rely on their records to carry out their mandate and to be accountable for their actions.[38]

Maintenance

Although the concept of "creation" as understood and used in InterPARES 2 is widely agreed upon in the archival community, that of "maintenance" may be less so. Many may see this term at first to be a synonym for preservation; however, although the two terms are related, they are in fact distinct.

Put in its most simple terms, record maintenance can be characterized as the actions performed on a record between its creation and its preservation. These actions may include description, storage, migration, reproduction, appraisal and provision of access, although not all of these actions are necessarily performed on all documents, nor in the order listed here. Also, once a given action is performed, it may be repeated an unlimited number of times, as needed.

Although the Terminology Database of InterPARES 2 does not define "maintenance," it does include two related entries, including "*Manage maintenance of kept records*,"[39] whose definition is "To provide overall control and co-ordination of the recordkeeping storage system and the records stored in the system by managing information about kept records and their digital components, placing the records in storage, maintaining the digital components and monitoring the performance of the storage system," and "*Maintain records in recordkeeping storage system*,"[40] whose definition is "To monitor the storage of kept records and their digital components and metadata, periodically back-up the recordkeeping storage system and, as necessary, correct problems with and update the digital components, and/or refresh storage

physical setting aside of the record in a recordkeeping system. For more information on this process, see the discussion of the Manage Making and Receipt of Records activity (A2.2) provided in the narrative to the Chain of Preservation Model that is presented in the Modeling Cross-domain Task Force Report. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_5_modeling_task_force.pdf.

[38] See InterPARES 2 Project Summary. Available at http://www.interpares.org/ip2/ip2_index.cfm.

[39] This phrase and its definition correspond to activity A3.2 in the Chain of Preservation Model.

[40] This phrase and its definition correspond to activity A3.2.3.3 in the Chain of Preservation Model.

media to ensure the records in the system remain accessible, legible and intelligible over time." These definitions outline three broad areas of maintenance: information about records, storage and "updating" records. Managing information about records essentially includes managing the capture, use and control of metadata about the records and the maintenance activities applied to them for the purpose of facilitating appraisal activities by the preserver and records indexing, storage, access and disposition activities by the creator. Managing storage includes overseeing the processes of placing the digital components of records and their metadata into storage (i.e., affixing them to digital media in the recordkeeping system), maintaining those components and metadata and monitoring the performance of the storage system. Lastly, "updating" records encompasses several related activities, including correcting problems with digital components in storage (i.e., dealing with stored digital components that cannot be located, retrieved, reconstituted or presented in accordance with current preservation strategies applicable to those records), updating the stored digital components (i.e., converting them via, for example, migration, standardization or transformation to persistent form, to ensure the records remain accessible, legible and intelligible over time) and refreshing the media on which the digital components are stored (i.e., copying or transferring the digital components from one digital medium to another, or otherwise ensuring that the storage medium remains sound), all of which involve careful consideration of various other maintenance-related issues, such as access restrictions, version control and the creation of an audit trail, among others.

Two key aspects of records maintenance can be found in the objectives of InterPARES 2.[41] One objective is "To identify and/or develop specifications for policy, metadata, and tools appropriate for the design of electronic infrastructures ensuring that…records are created accurate and reliable, and maintained and preserved authentic." Although this objective may seem at first to fall within the mandate of Domain 2, Authenticity, Accuracy and Reliability, it also has bearing on the work of Domain 1, since it deals with records creation and maintenance. While alluding to the record-making and recordkeeping systems ("infrastructure") in which creation and maintenance take place, this objective directly mentions metadata and other "tools" that can help maintain records. It also makes the distinction between maintenance and preservation, although it also shows that both actions are part of ensuring the authenticity of records.

A second InterPARES 2 objective is "To formulate methods for ensuring that…records are generated and maintained by the creator in a way that guarantees their accuracy, reliability and authenticity." Although this objective may also seem more appropriate to Domain 2, it nonetheless includes an important aspect of records maintenance: the fact that this maintenance is performed *by the creator*. This fundamental point is what essentially distinguishes maintenance from preservation and leads to the concept of the lifecycle of a record and how the notions of creation and maintenance fit within that concept.

Creation and maintenance in the context of the lifecycle of records

While the above discussion of "creation" provided the InterPARES 2 definition for "created record," the Terminology Database also includes a similar entry, that of "*record creation*." This term is defined as "The first phase of a record's lifecycle in which a record is made or received and then set aside for action or reference." This definition is essentially the same as the one provided for "created record," with the exception that it situates creation as taking place in the first phase of a record's lifecycle.

[41] InterPARES 2 Objectives. Available at http://www.interpares.org/ip2/ip2_objectives.cfm.

The notion of lifecycle was originally conceived in France, where it became known as "the three ages of documents." It classified records as current records, intermediate records or definitive archives, based essentially on where the records were kept: at their place of creation, an intermediate archive centre or in an archival depository. This concept is also sometimes expressed in the terms "active records," "semi-active records" and "inactive records." Later incarnations and versions of the concept of lifecycle shifted the basis of the definition from where the records were kept to what actions were performed on them. For example, in the theory of lifecycle that developed in the United States in the 1960s, there were two phases—the records management phase and the archival phase—each with four actions. Two of the actions in the records management phase were "creation or receipt" and "maintenance and use."

The InterPARES 2 definition of the lifecycle of a record is an extension of this last approach, although it transfers the notion of lifecycle from the activities that are carried out on records to the records themselves. The InterPARES definition divides a record's lifecycle into two phases, the first being when the records are still in the possession of the creator and the second when the records are in the possession of the preserver. What actions are performed on the records—and more importantly, who performs them—are key elements of the InterPARES notion of lifecycle.

If the actions are performed *by the creator* in the usual course of affairs for the purpose of those affairs and the creator keeps the outcome for further action or reference, then the actions are considered as maintenance and the resulting documents are considered to be the records of the creator. On the other hand, if the actions are performed *by the preserver* for the purposes of preservation and dissemination and not for the use of the creator, then it is clearly no longer a question of maintenance but of preservation, and the resulting documents are not considered to be the records of the creator but, rather, authentic copies of the creator's records.

As a result, it can be seen that the records lifecycle, as defined by InterPARES, "implies a shifting of responsibility for the records, from the creator to the preserver," and that the lifecycle is based in part on the use of the records, "and consequently on the purpose of the activities carried out on the records and on the person responsible for those activities, the creator or the preserver."[42] Moreover, as noted earlier, both creation and maintenance are actions that are carried out by the creator during the first of the two phases of a record's lifecycle. The two essential differences between maintenance and preservation, then, are the fact that preservation is carried out in the second phase of the lifecycle of a record by the preserver, not the creator, and for different purposes or uses.

Characterization of the Case Studies

This section presents the scope of the analysis relative to the records creators and their activities resulting in document creation. The answers to the Domain 1 research questions have been largely based on an analysis of the twenty-three case studies completed by the researchers of the InterPARES 2 Project.[43] In particular, the report focuses on answers given to the seven research questions pertaining to the Domain 1 concentration: Records Creation and Maintenance.

[42] Luciana Duranti (2005), "The Concept of the Records Life Cycle," PowerPoint presentation, slide 14 (unpublished).

[43] Of the original twenty-nine case studies proposed and approved for InterPARES 2, several were "retired," two remained uncompleted at the time this report was drafted (case study 22 - Electronic Café International: Aging Records from Technology-based Artistic Activities; case study 08 - Mars Global Surveyor Data Records in the Planetary Data System) and the four components of case study 09 (Digital Moving Images) were treated as four individual case studies, leaving twenty-three completed case studies that served as the basis of the present report.

The records creators

The following table presents the various types of creating bodies investigated by the case studies teams. The number of cases from the private and the public spheres are comparable. There are eight private organizations and eleven public ones. Besides these, four cases have a mixed structure. The ten artistic creators are mostly concentrated in the private sector; however, three are in the public sphere and two have a mixed status. The five scientific bodies are not concentrated in any one particular sphere. Among the seven case studies related to administrative governmental activities, five deal with administrative activities while the other two deal with private sector activities.

Table 2. The Case Studies' Creators

Status	Total	Type	Arts	Sci	Govt
Private	8	Individual	2	0	0
		Corporation	3	1	2
Public	11	Government	0	1	6
		Agency	2	1	0
		Cultural Center	1	0	0
Mixed	4	Partnerships	2	2	0

Six of the eight cases in the private sphere are artistic bodies. There are two individuals: a composer (case study 13, *Obsessed Again…*) and a performance artist (case study 02, Performance Artist Stelarc). There are three corporations: A small theatre group (case study 01, ArboCyber, théâtre(?)), a multimedia production company using industrial design methodologies (case study 09-1, Altair4 di Roma) and a large commercial movie-making company (case study 09-3, Commercial Film Studio). However, three cases are in the public sphere. There is a cultural center/public non-profit organization (case study 03, *HorizonZero*) and two agencies, one involved in film (case study 09-2, National Film Board of Canada) and the other in television (case study 09-4, WGBH Boston). Lastly, there are two bodies with mixed status that are, in fact, partnerships involving university laboratories (case study 10, *The Danube Exodus* and case study 15, *Waking Dream*).

The five scientific bodies are not concentrated in any one particular sphere. There is one case in the private sphere (case study 14, Archaeological Records in a Geographic Information System), two cases in the public sphere (case study 08, Mars Global Surveyor Data Records in the Planetary Data System and case study 19, Preservation and Authentication of Electronic Engineering and Manufacturing Records) and two university research groups with a mixed structure (case study 06, Cybercartographic Atlas of Antarctica and case study 26, MOST Satellite Mission).

Among the eight case studies related to governmental activities, six deal with administrative activities (case study 05, Archives of Ontario Web Exhibits; case study 18, Alsace-Moselle Land Registry; case study 17, New York Department of Motor Vehicles; case study 20, Revenue On-Line Service of Ireland; case study 21, Supreme Court of Singapore; and case study 24, City of Vancouver Geographic Information System). Two others deal with private sector activities: an Italian provincial body of a national cooperative (case study 25, Legacoop of Bologna Web Site) and an expert services company (case study 12, Antarctic Treaty Searchable Database).

Obviously, such a sampling could not comprehensively cover all possible activities in the given sectors. Nonetheless, the sample is, to varying degrees, representative of the various frameworks within which records creation is taking place in each sector of activity. The arts are mostly created in the private sphere on an individual basis or in informal or incorporated groups. Besides this model, there are also non-profit public organizations, particularly in more commercial sectors like film and television, where industry is more prevalent. Lastly, university laboratories are experimenting with the use of new media in the arts.

The scientific sample does not include a single individual, since researchers rarely work alone. They are instead often affiliated with organized research units based on partnerships between universities, private research institutes and government agencies. Their projects are funded by various public and private bodies and are under the leadership of a principal scientist accountable to the funding agency.

Public organizations are governmental by definition, but not all are dedicated to administration. Half of those studied are doing administrative or service activities. Besides these, there are two cases in which the private sector is closely related to government: a cooperative that is closer to a public structure than a corporation and an expert company led by two individuals, one of whom is a scholar. Lastly, it merits noting that public partnerships in the arts and science spheres and those with a mixed structure involved non-administrative but governmental concerns shared with university research groups.

Records creator contexts

By their nature, information technologies favour inter-connection and networking. Not surprisingly, therefore, the creator context reveals a collaborative dimension underpinning the process of records creation and maintenance. Although only four creators have a mixed structure, a second look at the other case studies reveals that many creators are actually working collaboratively in one way or another.

In the arts, an artwork may be created by an artist who is fulfilling a performer’s commission or working under contract on a given project. Collaboration may be between individuals—like a creative artist and a ballerina—or it may be between an individual and public or private bodies. In science, the works are largely based on a wide range of formal agreements with organized research units. In governmental activities, there are cases of joint ventures with private or public partners. This may come about when implementation is partly outsourced to a specialized corporation, or it may be for a given project involving two public bodies.

The collaboration context has the effect of distributing records creation among parties remote from each other—another dimension that new media favours. For instance, one case study associates an individual, a non-profit public agency, a private foundation and a project from a university centre. Each party created portions of the digital entity studied. The combination of all parts produced the artwork. The manager of the project was said to be from the university centre, which means that he could be considered as the party implementing the digital entity. However, the creative vision of the product was through the eyes of the individual artist and it was not clear which contributor was the originator of the project.

It is often the individual responsible for the creative vision who is the sole person to know how all the pieces fit together to interact and work. In large organizations, these instructions may be documented, but individuals often do not feel the need to do this, and the lack of such documentation may pose a problem. As one report states: “It is in the best interest of the composer to document as completely and accurately as possible his process of creation and the characteristics of each element of his work in order to facilitate future performances of his work, especially if the accurate reproduction of his intentions is important to him,"[44] because if "traditional instruments and tools generally remain available for use for long periods of time, technological elements are 'almost guaranteed to become obsolete within a very short period of time after the work's creation.'"[45]

Case study 13 (*Obsessed Again…*) illustrates this point. It involves the composition of a musical score for bassoon and interactive electronic work written in 1992. The equipment required to perform it is quickly becoming obsolete. To be performed using current technology, the composition will require recovery and a substantial reworking of both the interactive and electronic elements. It appears, however, that some components will not be updated due to software obsolescence and will thus need to be re-created. This process of re-coding and re-implementing aspects of the work may be seen not as *keeping* or maintaining the records, but rather as continuing the composition process (i.e., record-making) by *re-creating* many of the records. In such a case, it would be impossible to overcome obsolescence and maintain the authenticity of the documents in instances where the composer, who is the only individual intimately familiar with all aspects of the original performance, is no longer available.

Even though the underlying cases are different, a similar problem exists with larger, governmental structures. Case study 05, an archives Web exhibit project involving the Toronto City Archives and the Ontario Archives, reports that during the study, the development of Web sites

> …appeared to be an emerging business process in that the Web-based resources were being developed to fulfil a 'big idea' but there was no procedural context established in terms of which officers would fulfil which roles, or what records needed to be created and how they would be maintained. Rather, various individuals participated in the creation of the Web site on an as needed basis, sometimes through business activities that were already being undertaken but were now adapted or applied to the creation of the Web site. Each individual's involvement was 'trust-based.' For example, the scanning technician was not required to report on the settings chosen for scanning a particular item. The scanned component was used on a basis of trust—the scanning technician's judgement in the matter was neither recorded nor challenged.[46]

Obviously, some guidance is needed to make sure that "all the processes that contribute to the making and use of the same records will be explicitly documented." Although contributors to a project may document their own role in the creation of digital entities, there is often no agreement on what types of standards to follow or even a thorough understanding of what preservation of authenticity or reliability entails. Individual partners should be aware that they need a collaborative effort with consensus on how to preserve the work as a cohesive whole for future re-presentation or *re-creation*. In other words, collaborative partnerships must develop and have in place a strategy of preservation and dedicated responsibility for its application before beginning their creation activities, or as soon as possible afterwards.

[44] Jennifer Douglas (2006), "InterPARES 2 Project - Case Study 13 Domain 1 Research Questions: *Obsessed Again...*," 4. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs13_d1_questions.pdf.
[45] Ibid., 2.
[46] Jim Suderman et al. (2004), "InterPARES 2 Project - Case Study 05 Final Report: Archives of Ontario Web Exhibits," 7. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs05_final_report.pdf

Activities resulting in document creation

The activities of the arts focus cases studied can hardly be compared with traditional activity processes. Some aim to *create an artwork* and others aim to *support the creation of artwork*. In the first group, some are creating new media artwork in whole or in part and others are totally or partly Web products in nature. All the cases in the second group are carrying out recordkeeping activities for digital records contributing to or documenting artwork or performances.

Table 3. Activities Resulting in Document Creation in the Artistic Sector

Artistic activities	Web	DB	GIS	Files	Other
Creation of artwork	***4***	***2***	***0***	***3***	***3***
Virtual reconstitution of an ancient Roman archaeological site on DVD					X
Program files for an electro-acoustic music composition				X	
Commercial computer graphics animated films				X	
Multimedia performance art piece involving dance/movement, a soundtrack, live and pre-recorded video and remote-controlled interactions between performers and various digital and analogue devices	X			X	X
An "online magazine with three Web environments with an HTML portal, which serves as the entry point to all three environments"	X	X			
A "multimedia interactive piece that has taken two forms, a gallery installation which is no longer active and a Web site which is still live"	X	X			X
A Web site (*Ludosynthèse*) with a synthesis section on the past experimental theatrical performances of the troupe and an interactive section where users can re-create similar performances	X				
Support the creation of artwork	***1***	***2***	***0***	***3***	***0***
Stelarc's Web site used to support the conception and development of a performance artist who considers his own body as the primary record of his works	X			X	
A database system used to store and manage digital animation products and documentation related to production		X		X	
A system at a television station that operates in a mixed digital and analogue environment, but which has developed and is converting to a digital asset management system		X		X	

Table 4. Activities Resulting in Document Creation in the Scientific Sector

Scientific activities	Web	DB	GIS	Files	Other
Web-based science creation	***1***	***1***	***1***	***1***	
"Dynamic, interactive, Internet-based, open source atlas of the Antarctic continent for education, research and policy	X	X	X	X	

purposes”					
Support the science activities	***0***	***1***	***1***	***3***	
A database and a GIS to answer archaeological research questions		X	X		
Data captured by a satellite camera to support astronomical research				X	
Set of data from NASA’s Mars mission to allow further analysis and disseminate persistent research information				X	
Archival studies research to test a preservation process designed to archive persistent computer-aided design (CAD) records				X	

Like the artistic sphere, all five science focus case studies are producing born-digital entities. The purposes of these activities can be broken down into two types: one case is *creating Web-based science* and the other four aim to *support scientific activities* in a given field by providing reliable and accurate information to the scientific community.

Table 5. Activities Resulting in Document Creation in the Governmental Sector

Governmental activities	**Web**	**DB**	**GIS**	**Files**	**Other**
Recordkeeping	***3***	***4***	***0***	***4***	***0***
The computerization of a land registry in Alsace-Moselle (France)		X		X	
Web-based legal and financial transactions the New York State Department of Motor Vehicles	X	X		X	
Web-based filing of tax returns and tax payment in Ireland	X	X		X	
Web-based civil registry for the Supreme Court of Singapore	X	X		X	
Services	***4***	***2***	***1***	***2***	***0***
An American searchable database consisting of copies of Antarctic treaty and policy documents, used as an information resource	X	X			
A Web-based enterprise GIS system that enables data on municipal infrastructure and services in the City of Vancouver to be presented to the end user in the form of interactive maps	X	X	X	X	
A Web site that provides detailed and specialized information on a cooperative and its services, both to the general public and to cooperative members via a restricted area	X				
A Web-based project that created, promoted, publicized and managed archival outreach activities through the maintenance of a Canadian Archives Web site	X			X	

In contrast, the eight government focus case studies are mainly doing traditional activities applied to the digital environment. They use digital technology to *create and keep official records* or to *deliver services to citizens* in an interactive way.

None of the digital entities produced by these activities are easy to categorize. They are rarely just *one kind* of entity. For instance, the entity can be both a GIS and a database. However, it is possible to profile two types by considering their general form and features.

The first type involves hypermedia features. A majority of cases have such features. In the arts, they are mainly related to the creation of artwork activities; in the sciences, they are related to the Web-based science and data creation activities; while the majority of the government focus cases fall into this group. The literature makes the distinction between "closed" and "open" hypermedia features. The closed feature is defined as an autonomous digital entity in which links and bounds are internal. It has a Web format because of the nature of its language or software or because the creator wanted to enlarge his/her audience. It could also be a digital entity stored on a disk or a CD/DVD. Conversely, open features are found on Web sites that intrinsically function with external links in a network. They can be: (1) *interactive*: they use visitors' actions to create or change the work; (2) *generative*: they modify themselves according to the instructions of a program; or (3) *contributive*: they enlist the participation, voluntary or not, of visitors who may add material or simply react through e-mail, for example.

The second type of digital entity regroups digital components like files, program code, etc., that take place in a larger process of action involving other electronic and/or analogue means. It is exemplified by digital entities such as music files and software patches, computer graphic moving images, manufactured files and astronomical data.

Addressing the Research Questions

Question 1a

> *What types of documents are traditionally created (that is, made or received) and set aside in the course of these activities that are expected to be delivered online? For what purpose?*

As will be seen shortly, few, if any, of the case studies believe that they are doing traditional activities, although many of them do, to varying degrees, perform traditional activities applied to the digital environment. Instead, most claim to be—and are in fact—carrying out a new, non-traditional activity. They do not, therefore, believe that they are creating traditional types of documents in the course of these activities, since in their view the activities themselves are not traditional and therefore could not create traditional documents.

Despite this fact, from the various activities carried out in the three focuses, it is possible to extract or infer traditional types of documents that would be created in the course of these activities and then look at the types of digital documents that are currently being created to accomplish those same activities. This also makes it possible to determine if the purposes for which these documents are created have changed from the traditional to the electronic environments.

Artistic sector (Focus 1)

Despite the move to non-traditional hypermedia, performance art and telecollaboration, some aspects of the types of documents that are created in the course of artistic endeavours have remained the same. In fact, the core purposes for the creation of these documents and the types of documents themselves, in the abstract, have remained constant.

An example of this fact is provided by one of the uncompleted case studies in the arts focus, Electronic Café International, which exemplifies the types of documents that are traditionally created in the arts. The uncompleted Electronic Café International case study notes that, in general, for creators in the artistic field, documents are created to plan activities, execute and perform works/events, record portions of these and document and review these works/events. Despite a wide variety in the types of art produced (painting, sculpture, music, dance, theatre, photography, motion pictures, etc.), certain types of documents are created in a traditional artistic setting, including sketches, notes, film, photographs, sound recordings, musical scores, correspondence, contracts, reviews and news coverage.

Scientific sector (Focus 2)

Although increasingly relying on, and indeed inspiring, the advance in technology, science can also be seen as carrying out certain traditional activities and creating traditional documents. Any scientific endeavour involves the planning or design of a research experiment; the design, invention or modification of the proper apparatus to use; the collection of data; the pursuit of the research; the analysis of data; the communication of analyses; and the publishing of results.

As in the arts, the exact form and content of scientific documents vary considerably depending on the specific field or experiment, but it is possible once again to come up with a list of certain "typical" scientific documents. Documents that may be generated from the above activities include contracts, notes, sketches, diagrams, technical specifications, procedures or protocols, various collections or aggregations of data, analyses or transformations of these data, reports, correspondence, articles and presentations at conferences.

Governmental sector (Focus 3)

Government would seem to be the one of the three focuses in which the most traditional activities are being carried out, generating the most identifiable types of traditional documents. Again, in this focus there is a wide range of activities that may be carried out under the umbrella of "government." At its most basic, a government must provide essential services to its citizens, including registry and regulatory services; issuing licenses, permits and authorizations; providing information and access to the various branches of government; assessing and collecting taxes, fines and fees; devising and enforcing laws and regulations and maintaining civil status records.

A profusion of documents seems to be the hallmark or typical image of government. Some of the typical documents generated by government activities include laws and regulations, myriad forms and requests, various reports, minutes of meetings, correspondence, memos, notes, guides, announcements, civil status records or certificates, licenses, permits, receipts, passports or identity cards, election results, maps, plans, drawings and countless other documents.

Question 1b

> *What types of electronic documents are currently being created to accomplish those same activities? Have the purposes for which these documents are created changed?*

Many of the respondents in the case studies had difficulty responding to the previous question, as indicated, or believed that it did not apply in their case, because they believe that they are not engaged in a traditional activity and thus are not creating traditional documents. This equation of traditional activity and traditional documents—the former seen as necessarily producing the latter—was nearly universal among the case studies in the three focus groups. However, when examining the purposes for which these documents are created and the types of documents that are created, independent of form or media, it is clear that the differences are not so great between the traditional and electronic environments.

Artistic sector (Focus 1)

In the arts focus, respondents were convinced that they were carrying out a new activity that was non-traditional and innovative. As the respondent in case study 01 stated, “This question supposed a parallel between traditional and electronic activities, which is not the case for Arbo Cyber, théâtre (?). The *Ludosynthèse* is a new activity that is not making use of the creation processes used in traditional documents.”[47] Performance artist Stelarc goes one step further, actually integrating electronic components into his body for his performances. As such, his performances cannot be conceived without electronic means any more than they could be conceived without the artist himself. Two of the case studies, *The Danube Exodus* and *Obsessed Again...*, do not even discuss traditional documents in their final reports.

For the most part, from the point of view of technique, materials, form and presentation, the artists are indeed correct in thinking that they are doing something new and non-traditional. The performances or activities themselves are not performances or activities that are traditionally created and then converted to digital form, they are born-digital. However, the same types of performances have previously been done with traditional means and documentation. One of the case studies recognizes this fact. “Because *Waking Dream* is a unique creative work, it is not possible to speak of a traditional activity that is replaced or altered as a result of changes to technology; however, previous creative works of a similar type may have made use of analogue technologies where *Waking Dream* uses digital technology to record and save film, photographic and sound components of the work.”[48]

The activities of the creators in the ten arts focus case studies involve the integration of multimedia language in dynamic, interactive or collaborative artwork. Through the use of technology from the very beginning of the creation process, these activities are being carried out in a non-traditional manner and are related to new media art that creates born-digital products. The case studies can generally be broken down into two groups. In seven cases, the purpose is to *create* new media artwork in whole or in part. Four of the seven are totally or partly Web products in nature and can be defined as “Web Art,” characterized by their non-linearity, hypertextual and collaborative dimensions. The four case studies that are not actively creating new media artwork are *supporting the creation* of artwork by carrying out recordkeeping activities for digital records contributing to or documenting artwork or performances. Thus, they are creating “enabling records,” as defined earlier by Duranti and Thibodeau.

Despite the fact that the documents created in the course of these new media activities are born-digital and do not have the same form as traditional documents, their purposes are largely unchanged in the move from a traditional to an electronic environment. Performance artist Stelarc, who believes that his performances cannot be conceived without electronic means,

[47] Martine Cardin (2004), “InterPARES 2 Project - Case Study 01 Final Report: Arbo Cyber, théâter (?),” 48. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_report_english.pdf, and http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_report_french.pdf.
[48] Jennifer Douglas (2006), “InterPARES 2 Project - Case Study 15 Domain 1 Research Questions: *Waking Dream*,” 1. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs15_d1_questions.pdf.

admits that his documents are created "mainly for use in a digital environment,"[49] but that the essential purposes for document creation have not changed. At the Web-based *HorizonZero* magazine (case study 03), the same types of records would have been created as part of a traditional or analogue publication process and for the same purposes.

At WGBH Boston, where the creator did indeed change from a traditional to an electronic environment, this consistency of purpose in document creation is clear, despite the shift in the means and form of document creation. "The purpose for which these documents are created has not changed in switching from an analogue to digital management system."[50] However, it is Altair4 that most clearly states the fact that although the creation process has changed, the purpose has not. "The purposes of document creation have not changed. There has long been a search for increasingly sophisticated ways in which to represent nature. Thus, digitization is simply one step further in this process; it does not change the original intent of the representation impulse."[51]

So it can be seen that the digital environment, although changing the process and output, does not change the purpose of document creation in the arts. The traditional purposes are now simply being carried out in the digital environment. Past experience has shown that, traditionally, documents are created to plan activities, execute and perform works/events, record portions of these and document and review these works/events. As with Electronic Café International, digital documents are, like traditional documents, created during the processes of planning, implementing and performing works.

Examples can be drawn from the case studies in this focus. For planning purposes, Stelarc continues to use his documents for "the development of the performance [and] documentation of the execution of the performance."[52] In the category "execute and perform," at the National Film Board of Canada, "The records support either the creation or distribution of moving image materials."[53] Nearly all of the creators record portions of their work. Documents are also used to document or review a performance, notably at Arbo, where documents are seen "as witness to the past…[in which] they will form a new digital memory for the group."[54]

With many of the case studies in the arts focus, the role of the documents created is both retrospective and prospective. In other words, documents are created for the seemingly contrary purposes of bearing witness to past performances and guiding or informing future performances. They both document the original creation and provide a blueprint for re-creation. As noted earlier, Arbo's documents bear witness to the group's past, but this is not their only role. Documents "will be integrated to enrich the *Ludosynthèse*, either as witnesses to the past or as performance material to be used by spectator-users."[55] This second role is typical of the arts focus, in which part of the role of documents, whether traditional or digital, is prospective; that is, to provide instructions on how to execute or perform a given piece. For example, "as with traditional scores, the score for *Obsessed Again...* is created to provide a performer with

[49] Henry Daniel and Cara Payne (2004), "InterPARES 2 Project - Case Study 02 Final Report: Performance Artist Stelarc," 7. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs02_final_report.pdf.
[50] Geneviève Sheppard (2006), "InterPARES 2 Project - Case Study 09(4) Domain 1 Research Questions: Digital Moving Images - WGBH Boston," 1. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-4_d1_questions.pdf.
[51] Natalie Catto (2006), "InterPARES 2 Project - Case Study 09(1) Domain 1 Research Questions: Digital Moving Images - Altair4 di Roma," 1. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-1_d1_questions.pdf.
[52] Daniel and Payne, "Case Study 02 Final Report," op. cit., 7.
[53] Andrew Rodger (2006), "InterPARES 2 Project - Case Study 09(2) Final Report: Digital Moving Images - National Film Board of Canada," 3. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-2_final_report.pdf.
[54] Cardin, "Case Study 01 Final Report," op. cit., 53.
[55] Ibid.

instructions to perform the piece as it was intended to be performed. The additional digital entities serve the same purpose in that they describe how all of the components of the work…interact with each other to reproduce the performance."[56]

In the arts focus, although the purposes for document creation have not changed, the means of creation and the forms of the documents have. It appears that digital is becoming the new "traditional"—it is a new tradition in the making. And with this new tradition have come new uses or purposes of digital entities.

One of these new uses is the use of documents—often those documenting the performance itself—for promotion or publicity. The use of digital technology makes the integration and use of these documents much easier for artists, who can reach a wider audience through digital means. Although one of the documents created by the *Waking Dream* team is a Web site, which is used to promote the work, the potential of this avenue of publicity is fully realized by Stelarc, who continues to use his documents as a means to publicize his performances. The only difference is that the format of the Web site now allows for a wider publicity platform. Digital technology can bring a performance or its resulting documents to those who would otherwise not be able to see or hear about a given artist. This accessibility is not just due to distance, but also to disabilities, and digital technology can, to a certain extent, overcome these. In addition to publishing their Flash-based Web magazine, *HorizonZero* also provides "an accessible text-based site, published in both official languages,"[57] in keeping with W3C standards for accessibility.

Another new use of documents that is only possible, or at least greatly increased in the digital realm, is interaction. This can be interaction between the artwork and its spectators, as is the case with *The Danube Exodus*, where the digital entities are created "largely to display, and allow interaction with, the multimedia installation,"[58] or, in the case of *HorizonZero*, where the programming code documents allow for an greater degree of user/reader interaction than can be achieved with analogue publications. The interaction can also involve the various elements of the artwork or performance, as is the case with *Obsessed Again...*, where the performer, a microphone, an IVL pitch-to-MIDI converter, a Macintosh computer with MIDI interface, an external Proteus 1 synthesizer and a second amplification system interact with each other to produce the performance.

Lastly, another use of digital documents in the artistic field that is not possible with traditional documents is the creation of a virtual environment that is impossible in reality. This is the case in Arbo's *Ludosynthèse*, where pieces of past performances may be selected and re-combined in ways and in an environment that never existed and also in *Waking Dream*, where the team uses digital technology to help create the effect for the audience of having entered a waking dream state. These effects would not be possible without the use of digital technology.

Scientific sector (Focus 2)

The five science focus case studies are all related in one way or another to representing a physical or cultural phenomenon in space, be it geographical space or astronomical space. Like the artistic sphere, all are creating born-digital entities directly in an electronic environment that are not the result of the digitization of traditional activities. As observed in the case studies in this focus, the purposes of these activities can be broken down into two groups. One of the case

[56] Douglas, "Case Study 13 Domain 1 Research Questions," op. cit., 1.

[57] Brent Lee (2004), "InterPARES 2 Project - Case Study 03 Final Report: *HorizonZero*/Zero Horizon Online Magazine and Media Database," 2. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs03_final_report.pdf.

[58] Sally Hubbard (2006), "InterPARES 2 Project - Case Study 10 Final Report: *The Danube Exodus*," 5. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs10_final_report.pdf.

studies (case study 06, Cybercartographic Atlas of Antarctica) is actually *creating* Web-based science as a key deliverable of a geographic research project. The other four cases involve activities that aim to support scientific activities in a given field. This is to say, they handle or manage data to *support* research in a given field by providing reliable and accurate information to the scientific community.

Like the case studies in the arts focus, those in the science focus also see themselves as participating in a new activity, one that is non-traditional or in which the "tradition" was born in the digital environment. In the NASA case study, the activities observed are confined exclusively to the electronic environment, as may be expected from this creator. For the experiment conducted by the National Archives and Records Administration in the United States, "the digital entities pertaining to this case study are born digital as CAD records."[59] Again, the self-assessed "newness" by the creators in these case studies can be seen as stemming from the fact that the activities are conducted, and their supporting documents are created, entirely within an electronic environment that does not involve the conversion of traditional documents to digital ones.

However, like the case studies in the arts focus, the purpose of creating these documents is unchanged, since most documents are the by-products of the process of scientific research and are intended to provide information about the conclusion of an inquiry or to serve as the basis for further research or study. As exemplified by the NASA case study, documents are generally created to capture scientific data for further analysis and experimentation and for the dissemination of research information to the scientific community. The fact that the documents are born-digital does not change their purpose. The MOST project echoes this fact, noting that "The data in the digital entities will be used and analyzed in scientific publications and presentations to enhance our knowledge of stars."[60] Likewise, the documents created at the Center for Desert Archaeology by its Coalescent Communities GIS are created "to support archaeological research into the causes, tempo and spatial variability of the conspicuous population decline noted in prehistoric pan-Southwestern cultures of North America beginning circa A.D. 1300 and continuing into the early 15th century."[61] The purpose of the digital documents, therefore, is not new, even though some of the tools used to create the documents are. In other words, although scientific research can be innovative in its use of technology, and hence in its creation of digital documents, it, and the documents it creates, still serve a traditional purpose.

As was noted for the case studies in the arts focus, there are new uses that are particular to, or which result from, the digital technology used by the science focus case studies. Interactivity, seen in the arts focus, is a new purpose resulting from the use of technology that also benefits the scientific community. For example, in the progression of geographic representation from paper maps to their computerized counterparts to a more complex geographic information system (GIS), the entity in question "eventually [becomes] more interactive and expanded to include

[59] Geneviève Sheppard (2006), "InterPARES 2 Project - Case Study 19 Domain 1 Research Questions: Preservation and Authentication of Electronic Engineering and Manufacturing Records," 1. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs19_d1_questions.pdf.

[60] Bart Ballaux (2005), "InterPARES 2 Project - Case Study 26 Final Report: MOST Satellite Mission - Preservation of Space Telescope Data," 9. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs26_final_report.pdf.

[61] Richard Pearce-Moses, Erin O'Meara and Randy Preston (2004), "InterPARES 2 Project - Case Study 14 Final Report: Archaeological Records in a Geographical Information System: Research in the American Southwest," 19. Available at http://www.interpares.org/display_file.cfm?doc=ip2_c14_final_report.pdf.

multimedia,"[62] as is the case with the Cybercartographic Atlas of Antarctica, which uses this feature of the technology as a key component of its geomatic activity.

This last case study also demonstrates another new purpose that is derived from the use of digital technology in the scientific field: the creation of a "virtual" environment that does not exist in the real world. Again, this is a new purpose that the science focus case studies share with their artistic counterparts. The Cybercartographic Atlas of Antarctica makes extensive use of virtual reality fly-through and gaming technology to present information in a way that is not possible with traditional documents. In the case of the Center for Desert Archaeology, the use of digital technology allows for modeling the interaction between archaeological sites within the south western United States in a manner that, again, is impossible in a non-digital environment.

A final purpose of digital documents in the scientific field is one that was not observed in the arts focus; namely, the ability to increase the speed of performing an action, transaction or analysis, often expressed as the ability to accomplish tasks in "real time." In the realm of geomatics, the Cybercartographic Atlas of Antarctica demonstrates that "today, Web mapping involves generating maps from distributed data sets in real time."[63] As can be expected, this timesaving purpose of technology is also used by scientists in space research. The vast expanse of space is no longer a barrier to the real-time transmission and reception of data. At NASA, digital scientific data records are transmitted in real time through the Command and Data Handling Subsystem of the Deep Space Network to Mission Ground Control, where they are accumulated in a project database.[64]

Governmental sector (Focus 3)

The eight government focus case studies are mainly doing traditional activities that are applied to the electronic environment. They use digital technology to create and keep official records (e.g., case study 18, Computerization of Alsace-Moselle's Land Registry; case study 17, New York State Department of Motor Vehicles; case study 20, Revenue On-Line Service; and case study 21, Electronic Filing System of the Supreme Court of Singapore) or to deliver services to citizens in an interactive way (e.g., case study 05, Archives of Ontario Web Exhibits; CS12, Antarctic Treaty Searchable Database; case study 24, VanMap; and case study 25, Legacoop of Bologna Web Site). Despite the application of traditional activities to the digital environment, none of the digital entities created by these activities are easy to categorize, since they rarely fit into just one category. Nonetheless, it is possible to profile two types of digital entities by considering their general form and features.

The first type makes use of hypermedia features. Most of the case studies in the government focus have such features. In the arts, these features are mainly related to the creation of artwork and in science, to the Web-based science activities. The majority of the government focus case studies fall into this group. The second type of digital entity consists of groups of digital components such as files, program code, etc., that take place in a larger process or action involving other electronic and/or analogue means. These are related to digital entities like music files, software patches and computer graphic moving images in the arts and to manufacturing files and astronomical data in the science focus case studies.

---

[62] Tracey P. Lauriault and Yvette Hackett (2005), "InterPARES 2 Project - Case Study 06 Final Report: Cybercartographic Atlas of Antarctica," 26. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs06_final_report.zip.

[63] Lauriault and Hackett, "Case Study 06 Final Report," op. cit., 27.

[64] William Underwood (2005), "InterPARES 2 Project - Case Study 08 Final Report: Mars Global Surveyor Data Records in the Planetary Data System," 20–21. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs08_final_report.pdf.

As stated above, in the government focus there is a predominance of cases that are applying traditional activities of this field to an electronic environment. In these cases, the electronic environment seeks to mirror the traditional environment at the user level, as was seen in InterPARES 1. "Such records, although fixed digitally on relatively unstable media, are intended to approximate the physical documents generated in the course of established business procedures in well-understood judicial contexts."[65] The documents created as the by-products of these activities are specifically designed to appear to mirror the physical appearance of their traditional counterparts to provide a sense of familiarity, ease of use and comfort among the citizen-users. The visual similarity between the two types of documents (traditional and digital) is a key feature that was sought for the implementation of the electronic filing system for bankruptcy documents in the Singapore Supreme Court. Because the electronic filing system mimics the traditional paper-based one, "the records in the EFS mirror the paper-based system."[66]

Sometimes the traditional and electronic environments co-exist, at least temporarily, to facilitate the acceptance of, and transition to, the digital world. In Ireland, the Revenue On-Line Service (ROS) is an e-government application used to file tax forms and pay commensurate tax liability. Although ROS is used to replace paper-based transactions, the existing paper-based system is still available for users who prefer that format. A similar situation exists with the New York State Department of Motor Vehicles (DMV), whose On-line Services System is used to provide three core DMV business functions, including issuing and renewing or replacing drivers' licenses, vehicle registrations and titles of ownership. Users can also engage in related online transactions, such ordering a driver's record abstract or personalized license plates, paying fees and scheduling road driving tests. Procedures in the On-line Services System have been set up to mirror as much as possible the procedures for transactions conducted in person at a DMV office. Like Revenue's ROS, the DMV's On-line Services System is used to replace paper-based transactions, but the existing paper-based system is still available for users who prefer that format. Unlike the ROS situation, however, users of the DMV's online system must first establish a core record (i.e., a record created for each individual DMV customer containing information that uniquely identifies each customer, such as name, address, social security number and birth date), which can only be done during an in-person visit to a DMV office.

The implementation of an electronic system that appears to mirror a traditional one can be seen as *continuing* the tradition via electronic means. This is precisely the intent with the computerization of the Alsace-Moselle land registry in France. The region is very proud of its land registry tradition, which is unique in comparison to the rest of France,[67] and there is a feeling that the computerization project will ensure the continuation of this unique local land registry system in the future. If the appearance of the electronic system did not visually resemble the traditional one, citizens would not have the same level of confidence and would not make use of such a system, due to their unfamiliarity with the new system and the lack of intuitive and familiar features. If that were the case, the government's investment in the new system, as stated above, would be for naught.

---

[65] Domain 2 Task Force Report, 120.

[66] Elaine Goh (2005), "InterPARES 2 Project - Case Study 21 Final Report: The Electronic Filing System (EFS) of the Supreme Court of Singapore," 17. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs21_final_report.pdf.

[67] Jean-François Blanchette, François Banat-Berger and Geneviève Shepherd (2004), "InterPARES 2 Project - Case Study 18 Final Report: Computerization of Alsace-Moselle's Land Registry," 10. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs18_final_report.pdf.

As can be inferred from the conclusions of the arts and science focuses, and the desire in the government focus to mirror the traditional environment, the purposes of creating documents in the government focus have remained unchanged in the move from a paper-based to an electronic environment. In the above examples of ROS and Alsace-Moselle, the creators echoed their counterparts in the arts focus by stating that the purposes for which documents are created have not changed from those of a traditional environment. The only difference is the inclusion of one or more added purposes, as will be seen below.

The case study on the electronic filing system (EFS) of the Singapore Supreme Court is an excellent example of this aspect of the government focus. The purposes for which these documents are created have not changed. On the contrary, they have been augmented through the move to an e-government environment. The EFS has enabled the Supreme Court to facilitate the filing of court documents, to enable the quick retrieval of court documents, to improve access to records and information, and to manage and track cases, streamline workflow processes and improve case file security.[68] These points illustrate the purposes—both traditional and new—of document creation in the governmental sector.

First of all, documents in the government focus provide a service or information to citizens. This is the case regardless whether the activity is traditional or non-traditional. Documents created for the Alsace-Moselle land registry are created and set aside in compliance with French real estate law, "which dictates that the juridical status of a property…must be made publicly available to interested third parties by means of transcription within a land registry."[69] The Ontario Web exhibits are meant to inform the public. The Antarctic Treaty Database is used and is intended to be used as an information resource and the documents on the Legacoop of Bologna's Web site are created as a means of sharing information about the Legacoop's projects and to provide services to cooperative members.

Another purpose of documents in the government focus that has remained unchanged in the move from a traditional to a digital environment is that these documents are meant to provide citizens with access to their government. In Alsace-Moselle, the motivation for computerizing the land registry is to allow for remote access. Before the creation of the electronic database, no digital documents were created in the specific process of registry inscription and publication. Instead, individuals had to visit land registry offices in person to view an inscription. Singapore implemented their electronic filing system "to enhance access to justice and instil public trust and confidence in the court system."[70] Even in non-traditional activities, like VanMap and the Ontario Web exhibits, access is a major motivator in the creation of documents. The purpose of VanMap "is to provide the user with instant access to this information to support various functions of the civic government."[71] In the case of the Ontario Web exhibits, "Access is defined both in terms of access to unpublished or previously poorly described materials as well as in terms of remote and around-the-clock access."[72]

Like the other two focuses, in the government focus, these traditional purposes carried out by digital documents are augmented with new uses or purposes of e-government documents. As with the artistic and scientific sectors, providing or increasing interactivity is also a purpose of digital documents in the governmental sector. This point is best illustrated by VanMap, which

[68] Goh, "Case Study 21 Final Report," op. cit., 13.
[69] Blanchette et al., "Case Study 18 Final Report," op. cit., 2.
[70] Goh, "Case Study 21 Final Report," op. cit., 3.
[71] Evelyn McLellan (2005), "InterPARES 2 Project - Case Study 24 Final Report: City of Vancouver Geographic Information System (VanMap)," 6. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs24_final_report.pdf.
[72] Suderman et al., "Case Study 05 Final Report," op. cit., 13.

seeks to provide "an interactive, graphical representation of the data that allows the end user to see how the various features of the City relate to one another."[73]

A second new purpose of digital government documents, as was alluded to above, is to reduce the cost of a transaction while at the same time increasing the ease or flexibility of performing that transaction. For the Archives of Ontario, all of their intended purposes were "accomplished at less cost and with greater flexibility in a Web environment than in a physical one."[74] The technology used in the Antarctic Treaty Searchable Database allows for increased search capabilities, while the motivation for computerizing the Alsace-Moselle land registry was to allow for remote access, speedier processing times and increased storage capacity. Besides increasing the ease and flexibility of transactions, the use of new technology in the governmental sector also helps reduce errors and increase the accuracy of these transactions by reducing human intervention or interference. In Ireland, "the rationale for developing the online service is quite simple: '…people don't really want to see us and we don't really want to see them. The whole process should work without too much actual interaction. It should simply happen as a matter of course.'"[75] The system is also "consciously promoted as a means to reduce errors in tax returns. As the Commissioners had found that nearly 20% of all returns were inaccurate or contained human error."[76]

A final new purpose for transaction-based governmental services is particular to the government focus: the provision and assurance of electronic security. Biometric identification is used in the Alsace-Moselle system. In Ireland, the system is designed to maintain existing levels of confidentiality while incorporating a further level of security. "Revenue's requirement for a secure system dictated the use of PKI as an additional element."[77] The Singapore Supreme Court also makes use of digital certificates, which are generated in-house.

Question 2

> *What are the nature and the characteristics of the traditional process of document creation in each activity? Have they been altered by the use of digital technology and, if yes, how?*

In general, the traditional processes of document creation—taken in the abstract to mean the activities and steps involved from conception to creation—have not been discarded in the move from a traditional to a digital environment. In some cases, the process is seen as continuing the tradition, but with electronic means. Technology has allowed the creator to carry out a greater portion of the creative process him or herself, which increases the ad hoc or individualistic nature of creation in fields where this is already the case. The main change in the creation process is an increase in the speed with which the process is accomplished and the inclusion of additional steps for verification or to take into account certain features or limitations of the technology used.

---

[73] McLellan, "Case Study 24 Final Report," op. cit.
[74] Suderman et al., "Case Study 05 Final Report," op. cit., 14.
[75] John McDonough, Ken Hannigan and Tom Quinlan (2005), "InterPARES 2 Project - Case Study 20 Final Report: Revenue On-Line Service (ROS)," 1. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs20_final_report.pdf.
[76] Ibid., 2.
[77] Ibid., 70.

Artistic sector (Focus 1)

In the arts focus, the activities and processes impacting the creation of documents obviously involve the development and implementation of an artwork or performance. Despite the wide variety of forms of artistic expression, the development of an artwork or performance, independent of the media or artistic field, involves some or all of the following document-generating activities: proposal and contract writing; applications for funding or other correspondence; research; and developing and perfecting notes, sketches and/or instructions for elements of the overall artistic design and the implementation of a performance, artwork or installation.

In the digital environment, documents are created to reflect these same traditional functions. The functions have not changed, although the means to achieve them—the activities—have. While some documents are created by digitizing traditional documents, including photographs, audio clips, technical drawings and some text documents, other documents, most notably text documents resulting from the creation process, are born-digital. More and more elements and products of the artistic process are also being born-digital, as exemplified by digital photography, animation and music.

The arts focus case studies reflect the unchanging nature of the process of document creation in the move from a traditional to a digital environment. In the case of Altair4, in archaeological films, paper drawings and watercolours would traditionally have been used to recreate an archaeological site. Now, the watercolours and paper drawings have simply changed to pixels. At the National Film Board of Canada, a number of activities traditionally have characterized—and still do characterize—the creation of animated films and result in a variety of documents. The majority of these documents are now either born-digital or scanned and brought into the digital domain. Some artistic activities, such as performance art, involve purely born-digital documents created in the digital domain that are not the result of changing from a traditional to a digital environment. The *Waking Dream* project claims that there is no traditional activity that the activity under investigation replaces, nor is there a traditional process of document creation that is replaced. However, performance pieces have been created prior to *Waking Dream*.

What is being witnessed in the arts focus case studies is the *continuation* of the artistic tradition in the digital environment. The processes are largely the same, based on the long-established artistic principles of each field. It is not the process that has changed, but the environment in which this process is carried out and the specific actions used to carry it out. The tools and materials used in creation are now digital, not analogue. Once again, Altair4 provides an excellent example of this point. The basic methodology followed in the geometric representation that it carried out in a digital environment is based in rules established during the Renaissance. The process of document creation used to model the components of the villa is governed by the archaeological practice of proceeding from the front to the interior. This did not change because a computer was used instead of paints and brushes. In the case of the commercial film studio, the process of production has not changed, in that once an individual receives an entity, whether digital or non-digital, he or she completes the required artwork or manipulation and passes the work on to the next person in the production process.

The process of document creation in the arts focus is largely variable or ad hoc. It is either an individual process or similar to one in its subjectivity or lack of formality. In the case of the National Film Board projects, it is not possible to generalize a workflow or document creating process, because each animator follows his or her own steps to create the final product. The collaborative effort to create *The Danube Exodus* was not collaborative in the sense that all parties worked together throughout the process. Instead, each of the groups involved in the creation of the installation developed its own creative process and then brought the results to the others. These processes were described as flexible and capable of adapting to circumstances as they arise. In other words, they were not fixed or formalized. The *Waking Dream* project also involved an ad hoc creation process. No formal procedures were followed during the creation of the performance. The three co-authors conceived of the idea together and then divided tasks according to each person's creative talent and area of expertise. The result was the collaborative work of all partners, but the partners did not work together; each developed his or her own process to arrive at a collaborative conclusion.

This individualistic, ad hoc nature of the creation process is not exclusive to the digital environment. However, it is easy to see why such a situation is continued in—even thrives in—a digital environment. The *HorizonZero* case study points out that, whereas traditional environments—especially "institutional" creators such as magazines or film studios—may have pre-established production mechanisms in place, new media producers often lack the managerial background to implement these mechanisms. This difference is attributed to several factors.

First, the environment in which new media objects are created is very different than that in which traditional media objects are created. The contributing creators of new media objects are often remote from each other and from the production environment. Also, new media productions frequently lack the financial resources to implement and monitor effective managerial practices. The Internet is often used as a means to reduce cost, although not necessarily to save money that the production has, but as a means of coping with the fact that the production has little or no money to begin with. Lastly, new media productions usually employ fewer people than traditional media productions, so that individuals often create and manage the documents for which they are responsible, seemingly reducing the need for strict document management practices.

Although the creation process is largely the same for traditional and digital environments in the artistic field, there are key differences between working in a digital new media environment and a more traditional print environment. These are not so much changes to the creation process, but additions to or new aspects of the process.

A digital environment generally provides an artist with more flexibility or freedom in the creation process. For example, traditional print and film media are largely fixed, while digital media has a more flexible nature. Besides adding flexibility to the creation process, this fact also leads to an increased need for procedures to be put in place to determine when a digital media object is considered complete and should not be altered. In the case of animated film, traditional animation was tightly controlled in the sense that film and frame size, frame ratios and projection speed were standardized. With digital animation, the frame ratio can be whatever the animator wishes it to be and the frame rate depends only on the video system being used. The WGBH case study also showed that with the digital asset management (DAM) system, production teams are now able to create their own footage logs during the actual production process and not months after, as with a traditional system. Also, the DAM system allows for the added capability of creating storyboards using film clips and re-purposing the footage for other media types. Thus, although the basic creation process has not been fundamentally changed, certain new, flexible features have been added to it.

Another novelty in the creation process introduced by the digital environment is increased interaction in artwork or performances. Altair4 has shown that the main element that has been altered by digital technology is the fact that the end result allows the user to interact with the artistic representation in a new way that would not be possible in a traditional environment. For the musical score of *Obsessed Again...*, where the traditional process may have included testing and modification of the score following rehearsal with instruments, the creation of electro-acoustic music includes the "virtual" development of many of the interactive elements that will ultimately make up the final work.

Lastly, the creation process in the artistic field has been adapted in some cases to take into account the technology used or because of the possibilities of or limits to the digital technology. The creators of *Obsessed Again...* felt that the traditional compositional process had been modified by the addition of new steps to accommodate the use of digital technologies. In the case of the uncompleted Electronic Café International case study, digital technology affected document creation in the sense that ECI's works or performances have been *limited* by available technologies and their functionality.

Scientific sector (Focus 2)

In the science focus, there is only one case that exhibits a variable, subjective creation process similar to the arts focus. For the Coalescent Communities database of the Center for Desert Archaeology, both the system and the variety of activities carried out through the GIS display an ad hoc nature in regards to the creation process. Like the commissioning of an artwork, the use of the GIS is usually dictated by a specific question or project, so a specific problem is addressed and a project file created for each project according to a process that is more or less specific to that project.

However, for the rest of the science focus case studies, document creation takes place in a much more formalized and controlled environment, with pre-determined processes including the collection, analysis and preservation or communication of data. Most of the case studies take place entirely in a digital environment with a born-digital creation process that is often wholly or partially automated through the use of technology. As with the NASA case study, in which standard data products, documentation, index tables and archive volumes are generated automatically, the majority of digital entities in the MOST project are also the result of automated processes of data gathering, packaging and reduction (analysis).

It is the desire and possibility of translating the collected data into a neutral or open source format that characterizes the majority of case studies in the science focus. In the NARA experiment, the original digital entities are altered by the use of digital technology in the interest of preserving these records and ensuring their authenticity, reliability and usability over time through persistent object preservation. In the case of the Cybercartographic Atlas of Antarctica, the increased use of multimedia objects led to the development of interoperability standards, open source specifications and more metadata elements for both information objects and their relationships.

One of the characteristics of the creation process for science focus case studies in the digital environment is the inclusion of elements that permit increased interaction, which is only available through digital technology. The increased use of multimedia objects has lead to increased interactivity, which is especially evident in the case studies involving geomatics and the use of GIS, such as the Center for Desert Archaeology and the Cybercartographic Atlas of Antarctica. That is because, today, Web mapping can involve generating maps from interoperable, distributed datasets in real time. This fact calls for increased interactivity between the user and the system to create these real-time results and also an increased use of multimedia objects that make up the presentation of the datasets to the user. As the field progresses, this type of interaction is leading to the development of the interoperability standards and open source specifications referred to above.

Another use of technology in the science focus, as stated above, is the automation of the creation process and the reduction of human input. This phenomenon has come about in an effort to reduce errors through the automated verification of data and can be reflected in the desire of many creators to increase the accuracy of their data from the very beginning of the creation process.[78] In this use of technology, one can see the implementation of the concept of the system as an agent of the creator with no physical, real-time involvement of the creator itself, as discussed in the InterPARES 2 definition of record.[79] Uses of this technology include the validation of data archive volumes by NASA and the automatic creation of files by the various software programs in the MOST system, with a lack of "human involvement."[80]

Although science largely makes use of the benefits of technology in the creation process, it is sometimes confronted by the limits of this technology, which therefore also limit or otherwise affect the creation process. A notable example from the case studies is NASA, where technology has affected document creation in the sense that NASA is limited by the current capabilities of the technology and instruments used. These limitations not only concern which data can be collected, but also the accuracy of those data, since accuracy is highly associated with both the method of data collection and the type of instrument used in the collection. Nonetheless, it is often exactly such limitations that lead to the development of new technology and new standards.

Governmental sector (Focus 3)

Most of the case studies in the government focus deal with a traditional activity being carried out in a new way. Therefore, the process of document creation is largely the same as for the traditional environment; it is simply transposed into the digital environment with the possible addition of certain steps in the process to take the technology into account. For the Alsace-Moselle land registry, the process is not changed; it is simply automated with the use of technology, as is the case for the electronic filing system of the Singapore Supreme Court, the Revenue On-Line Service of Ireland and the On-line Services System of the New York State Department of Motor Vehicles.

While the above-mentioned systems are examples of traditional registry-type activities that are now being carried out online, the lack of change in the creation processes is not limited to what can be seen as traditional government activities. At the Italian cooperative, the Legacoop of Bologna, an increasing number of documents are born-digital, although they mirror traditional documents in appearance and intent. The group's newsletter and job postings are simply published online, as opposed to in print form. The underlying process has not changed. Even in the case of the Ontario Web exhibits, the creation process is not something that is entirely new or different. Although creating a Web-based exhibit was seen as taking place in a nascent creation context and was described as an emerging business process, this nascent activity should be nuanced, however. The creator acknowledged that although the creation of Web exhibits is an emerging business activity for it, the creation of exhibits is not a new activity for archival

[78] Although the term "data quality" is often used in the scientific field to include the notion of accuracy as understood by InterPARES, the term "accuracy" is used here to be consistent with InterPARES terminology. Further discussion of issues surrounding the concept of "data quality" can be found in the Domain 2 Task Force Report in the section titled "Conceptual analysis: authenticity, accuracy and reliability in the literature of the sciences."
[79] Duranti and Thibodeau, "The Concept of Record," op. cit., 7.
[80] Ballaux, "Case Study 26 Final Report," op. cit., 10.

institutions in general and the creator has previously been involved in creating traditional exhibits.

There are, of course, some exceptions to the observation that most government focus case studies have similar creation processes for the traditional and digital environments. These exceptions, however, should be seen as nuances to the above affirmations that are brought about by pursuing non-traditional–type activities. Due to the creator's perception that the activity in question is a non-traditional or nascent business practice, some of the otherwise strict control mechanisms in the creation process were slacked or ignored in the digital environment.

In the case of the Ontario Web exhibits, in both institutions, the business process leading to the creation of Web exhibits was in a formative stage. In each organization, some aspects of the process were clearly defined, whereas others appeared to vary or be ad hoc in nature. Also, because the activity of creating Web-based exhibits was seen as an emerging business process, there were no formal record-making practices in place. Various individuals participated in the creation of the Web site on an as-needed basis, sometimes through business activities that were already being undertaken but were now adapted or applied to the creation of Web site exhibits. Similarly, at the Legacoop of Bologna, there were no specific criteria or controls over the creation of digital documents, as opposed to their traditional counterparts. Despite the fact, as mentioned above, that these activities are normal activities of the creator previously carried out in a traditional environment that were being applied to a digital environment, the "newness" of the digital world partially blinded the creator to the fact that the same creation processes could be carried out with slight additions or modifications. As seen above, both of the two creators carried out some aspects of their traditional creation process, while other aspects were ignored or simply not applied in the digital environment.

As with the science focus, some of the case studies in the government focus used technology in the creation process to reduce or eliminate "human interference" as a means of reducing errors through system automation. In the Alsace-Moselle land registry, the use of digital technology automates and expedites records creation. Requests for inscription are received electronically, using custom software that connects to the land registry database to retrieve the information relative to the property in question. Once the request is received, it is dated and a digital file is created containing all associated documents as scanned image files. An ordinance project is prepared automatically and is transferred to a judge's in-box. The judge then "intervenes" by verifying the information and electronically signing the ordinance project. An ordinance is then created by the system, and the relevant fields in the database are updated automatically. In the traditional process of document creation in Ireland's Revenue On-Line Service, Revenue employees manually enter information from analogue forms into databases. In the electronic system, certain fields of the tax forms are pre-populated and automatically verified to reduce the number of errors that were apparent in analogue formats. Information is also now added to the database automatically, rather than being typed in manually by Revenue employees. In other systems where automatic pre-population of forms is not possible, there are still control mechanisms in place to reduce human error. In the electronic filing system of the Singapore Supreme Court, law firms are required to enter information under a prescribed documentary template in EFS before submitting it to the courts. Any deviance from the template is rejected, which helps automate the process and reduce differences and errors.

A final aspect of document creation in the government focus that is the result of the move from a traditional to a digital environment is the inclusion of security measures as a means of assuring authenticity from the very beginning of the creation process. In Alsace-Moselle, biometric technology is used. The judge identifies him or herself through a fingerprint scan and a smart card containing his or her private signature key and electronically signs the ordinance project. Based on the need for a secure online environment, Ireland's Revenue On-Line Service (ROS) is also regulated using private key infrastructure. Document creation is done, therefore, in a controlled environment. ROS users are required to obtain access numbers: individuals require an ROS Access Number and tax agents require a Tax Agent Identification Number to use the system. In Singapore, the Supreme Court itself manages the public key infrastructure process that results in the issuance of a smart card to only those solicitors who possess valid practicing certificates.

Question 3

> *What are the formal elements and attributes of the documents generated by these processes in both a traditional and a digital environment? What is the function of each element and the significance of each attribute? Specifically, what is the manifestation of authorship in the records of each activity and its implications for the exercise of intellectual property rights and the attribution of responsibilities?*

There was considerable variability in the awareness and use of document elements and attributes in the three focus areas. Perhaps this variability can be attributed to the differences in the creation environments and the legal and professional requirements present in each focus. A general progression of the understanding and use of elements, attributes and the manifestation of authorship can be seen from the arts focus to the science focus, finding its most clear understanding and systemized use in the government focus.

Artistic sector (Focus 1)

Most of the case studies in the arts focus had no idea of, or at least no concern for, the "elements" or "attributes" of their documents, as these terms are defined by InterPARES. As is understandable, aesthetics are prioritized in any artistic endeavour and the creators often pay no specific attention to the attributes of digital entities. As a result, few formal elements were identified in the records. For example, at Arbo the *Ludosynthèse* records contain several elements, but these are not standardized and the digital records' details of date, time and place are not always noted. What is more, subjects are not explicitly inscribed on the digital records.

When pressed about the matter by the Project's researchers, many of the case studies' creators spoke of "elements" in terms of the file type or format in which a document is manifested, and it became apparent that this term is widely understood as such in the artistic community. Sometimes, responses were given that listed the specific file types created or used by the creators, or instead echoed Stelarc's response that "key elements of the digital entities on the artist's Web site are text, still and moving images and sound."[81] In some instances, the case studies do not describe the formal elements and attributes of the documents generated during the creation processes, except to describe the hardware and software used by each of the creators. In some cases, such as *Waking Dream*, there is either limited information in the case study final report about the formal elements and attributes of the various digital entities, or the formal elements and attributes of the creator's documents are not noted at all.

[81] Daniel and Payne, "Case Study 02 Final Report," op. cit., 7.

When they are identified in the arts focus case studies, elements and attributes are largely defined or limited by software specifications. It seems that in the arts there is little or no intentional or even conscious capture or notation of elements by the creator. Many creators in this focus used off-the-shelf graphics or production software, with little or no modifications. At Arbo, although technological constraints certainly influence the form of certain records, these are not seen as affecting the record's function. As a result, the formal elements and attributes of the documents are determined by the specifications of the individual software programs used in the creation process, whether or not the creator is aware that the software is capturing the documents' elements and attributes.

The use of off-the-shelf proprietary software not only limits the capture of elements, but also access to the documents themselves. Because these documents are created using proprietary software, they can generally only be accessed using that software and, in some cases, the particular operating system. For example, the computer code used to read the remote control dowser in *Waking Dream* is written in a version of Visual Basic Project Manager developed to run on a Windows 98 platform. The code can only run in Windows 98 because it requires access to functions that have been disabled in subsequent Windows operating systems. Also, the PowerPoint file currently only works on a Macintosh computer running OS9. As such, the very elements that make up the record are confined to the software and operating system in which they were created. Thus, they are not interoperable and run the risk of becoming obsolete.

Authorship is a concern for artists, but it is often not formally manifested in works created by individuals or in loose partnerships or collaborations. Sometimes, authorship is only attributed when works are collected or published, not at the moment of creation. This is the case with Arbo, which is creating a "final credits" page for the *Ludosynthèse* on its Web site. This page will identify those who worked on the *Ludosynthèse* and the collaborators, photographers and others who contributed prior to digitization. The chronological section and the "Documentation" file will also introduce each participant in original performances. The digital records are therefore not signed, except to identify those who participated in the original analogue production. In this case, the signature is visible to users and is usually included to ensure authorship rights. Signatures are not, however, necessarily attached to the records, as they can be located on the credits page only.

It is often assumed by artists that simply keeping their works in their own possession or allowing controlled access to them is enough to prove or ensure authorship. For example, with *Obsessed Again...*, the composer is the author and sole possessor of all digital entities created during the composition of the piece. External users do not have access to the master copies of the digital elements used to create the piece. Although the composer's authorship is protected under copyright legislation, it is not clear from the final report how authorship is manifested in the individual records themselves, if it is at all. With individuals, possession or controlled access to their works is largely put into practice by keeping their documents on their personal computer or by the presence of the documents on a proprietary Web site. This fact is best demonstrated by Stelarc, who is considered to be the author of all the documents on his Web site, as it is he who chooses what to include and post.

However, authorship issues become muddled in the cases of partnerships or collaborations. In *The Danube Exodus*, it is difficult to ascertain exactly how authorship is manifested in the documents, but it is clear from the interim case study report that authorship is an issue in this project, since different components of the installation are created by different groups and individuals. The question of authorship is also interesting in the case of *Waking Dream*. The creation team consists of three people and each is responsible for the creation of certain components. However, authorship has not been made manifest in any of the formal attributes of any of the digital components.

The larger and more "official" the creator in the arts focus, the clearer the issue of authorship becomes. Larger and more formalized groups usually have some clear means of indicating authorship of their records. At Altair4, the manifestation of authorship in the records of each activity and its implications for the exercise of intellectual property rights and the attribution of responsibilities are all decided by the three heads of the company. At the online publication *HorizonZero*, media assets that are saved to the ZeroHorizon database, including assets designed and developed by individual contributors, are tagged with *HorizonZero* metadata in accordance with CanCore standards, including those noting authorship.

Some more organized groups or collaborations in the arts focus have written contracts or even a rights management system specifying authors' rights and intellectual property issues. At *HorizonZero*, individual contributors and artists retain copyright over their work but waive moral rights, thereby permitting *HorizonZero* and the Banff Centre to reproduce the work in print or digital form in perpetuity. In other words, although individual artists and writers are the authors of their work, many of the rights associated with authorship belong to *HorizonZero* and the Banff Centre. Specifically, *HorizonZero* is allowed to use an artist's work for any purpose without paying royalties. In the uncompleted Electronic Café International case study, ECI is the author of the digital documents in the legal sense. In particular, Sherrie Rabinowitz and Kit Galloway hold intellectual rights over entire projects, as they envision, develop and manage the collaborative works. More "industrial" creators, such as WGBH and the National Film Board of Canada, have electronic rights management systems that provide information about rights pertaining to a given production and determine the royalties that must be paid out when a production is sold or broadcast.

Scientific sector (Focus 2)

For larger creators in the science focus, elements are often formalized or structured by professional standards or practices, or, in some cases, by well-documented in-house standards and practices that are specific to a project. At NASA, digital entities take the form of structure objects within the Planetary Database (PDB), which outlines the format in which the scientific data appear in PDB labels. Standards for the form and description of elements are documented in the PDS Standards Reference, which provides a detailed description of each label and the Planetary Science Data Dictionary, which, in turn, provides definitions for all attribute names used in resource descriptions. PDS data also adhere to nomenclature standards, which define rules for constructing Data Element and Data Object names. In the MOST project, there is an internal document that describes the descriptive fields of the FITS (i.e., Flexible Image Transport System—a standard astronomical data format endorsed by NASA) files. The digital entities are also given a unique, standardized name, based on the target (i.e., the star) and the time.

Larger, more structured creators also often use metadata or professional standards to reference elements. For the Cybercartographic Atlas of Antarctica, the data are fully referenced within modules or within the metadata or are embedded within the digital objects. Remote access to some data is possible on the fly when a map is created, in keeping with the Open Geospatial Data Consortium Standards. At MOST, the metadata schema that is used was created by the project researchers and is specific for the data files created in the MOST project. It is based on experience and best practice in the astronomical community and on the foreseeable use of the records in the future. Some of these metadata or descriptive fields in the FITS files are mandatory because of the file format.

As with the case studies in the arts focus, for small corporations and individuals in the science focus, document elements are not formalized, although some practices may be used informally, based on perceived best practice, practicality or functionality. At the Center for Desert Archaeology, the process for creating and maintaining digital entities is ad hoc and lacks systematization, mainly because of financial and time constraints. Also, the procedural context is not rigid or always predetermined, due to the small number of people creating the GIS and the fact that it is a work of a non-governmental organization. Within the ArcView software, the user could create, manage and edit metadata based on accepted standards in the field. However, the nature of the archaeological site source information used is based on an idiosyncratic, in-house decision rather than on any established metadata standard. Nevertheless, this practice is seen as improving the reliability of the database as a trusted source of archaeological information.

Authorship is clearly manifested in the documents of larger creators in the science focus. The individual author is sometimes noted, but in most cases the sponsoring body or collective creator is often attributed authorship over the individual scientists, who are mostly regarded as producing a work-for-hire. This is a notable difference from the case studies in the arts focus, as exemplified by *HorizonZero*, where individual authors are noted, although the collective creator retains certain intellectual rights. In the Canadian copyright legislation that applies in this case, rights to a work are divided into financial rights (article 3) and moral rights (articles 14.1 and 28.2(1)). Although financial rights may be sold, granted or waived, moral rights, such as acknowledgement as the author of a work, cannot be waived or ceded.

For the Cybercartographic Atlas of Antarctica, individual module content creators are identified when their content is discussed in various academic papers, presentations and reports and on the project's Communication Web site. Authorship and responsibilities for the Atlas itself are attributed to the Geomatics and Cartographic Research Centre. In the case of the Coalescent Communities database at the Center for Desert Archaeology, the creator and author is the Center for Desert Archaeology, but the name of the originator is the GIS Specialist, who is also the writer. Within the archaeological site entry form, there is a special sign in the form of the organizational logo. The creator views the special sign as one that denotes authorship and intellectual ownership of the dataset as a whole; that is, as a unique amalgamation of individual datasets, each of which possesses its own authorship and intellectual ownership attributes. At NASA, the agency is the author of all documents created in the Mars Global Surveyor Mission. For smaller groups, authorship is not as much a concern and as a result is not as formally noted.

Similar to the arts focus, in which certain intellectual property rights were ceded to a publishing authority to reproduce the work, in the science focus intellectual property issues are sometimes limited to a proprietary period or are simply waived or reduced to share data with the scientific community. This is in keeping with the scientific method, which results in data being shared with other researchers and the public in the interest of furthering science. However, although the data are shared, certain moral rights, such as paternity, are retained by the original creator. An example of this situation is the Cybercartographic Atlas of Antarctica, in which much of the data used in the creation of the atlas—a non-commercial research product—can be used at no cost as part of the Antarctic Treaty System. The product nonetheless includes typical intellectual property issues such as license agreements, use rights to objects and data and copyright.

Governmental sector (Focus 3)

By and large, in the government focus there is a greater awareness of the formal elements and attributes of documents than that which is found in the other two focuses. Metadata are often used to define or describe the various elements, similar to the science focus, and may be automatically or intentionally generated, depending on the creator and the software used.

Often, elements were correctly understood by creators to be the intrinsic and extrinsic elements of their documents. Perhaps the most easily and consistently identified elements were those related to Web site documents. Elements and attributes that are considered integral to the validity and completeness of such documents (intrinsic elements) include: navigation links, the creator's logo or visual identity signs, a privacy policy, terms and conditions of use, site content and copyright statements. Elements that constitute the material make-up of the document and its external appearance (extrinsic elements) include a Web page template, cascading style sheets, navigation bars and a feedback form.

Many elements are designed for the purposes of providing and assuring security, privacy and authentication. Formal elements and attributes include digital certificates and signatures, annotations and electronic seals. Digital certificate and signature elements include the name of the certificate owner, the dates of issue and expiration and the user's public key and unique login number and a date/time stamp.

As a rule, authorship is formally manifested in the government focus, mostly through logos or other visual cues or by means of a formal statement of copyright. Often, the URL of a Web site is seen as proof of authorship, as an indication of the corporate domain. For example, in the URL http://www.archives.gov.on.ca/english/exhibits/index.html, "gov.on.ca" indicates that the site resides in the domain of the Government of Ontario, Canada. In the Revenue On-Line Service of Ireland, Revenue's authorship of its documents is reflected in the consistent URL naming, which provides the ROS with strong ties to Revenue.

A particularity of the government focus is that the authorship of, or responsibility for, the records is sometimes separated from the responsibility for maintaining the digital environment. For example, in the Alsace-Moselle land registry, the judge remains the author of the ordinance and the information in the computerized registry is based on the ordinance. However, the GILFAM (i.e., Groupement pour l'Informatisation du Livre Foncier d'Alsace-Moselle—the administrative body specifically charged with computerizing the land registry) is responsible for the maintenance of the database system. In the case of the Legacoop of Bologna, although the content of its Web site is created in-house, the Web site is technically managed by an external Web agency that is responsible for ensuring infrastructural services, posting data and developing and maintaining the technical and graphic aspects of the site. Other creators, such as the Irish Revenue and the Singapore Supreme Court, maintain their own systems in-house.

As has previously been seen with the types of documents created in traditional and digital systems, in the government focus, the physical elements of documents created in the course of more traditional registry or service activities often appear to replicate the appearance of paper-based elements. In the computerized land registry system of Alsace-Moselle, form and function (physical elements and their behaviour) remain largely identical in the traditional and digital environments. In the digital environment of the Revenue On-Line Service of Ireland, it should be noted that the use of the PKI environment in conjunction with digital certificates is analogous to an individual using his or her PPS (Personal Public Service) number and signature in the analogue environment. Tax forms and their elements are designed to appear visually consistent with existing paper-based forms, but have added levels of pre-population and dynamically generated content.

Question 4

> *Does the definition of a record adopted by InterPARES 1 apply to all or part of the documents generated by these processes? If yes, given the different manifestations of the record's nature in such documents, how do we recognize and demonstrate the necessary components that the definition identifies? If not, is it possible to change the definition maintaining theoretical consistency in the identification of documents as records across the spectrum of human activities? In other words, should we be looking at factors that make a document a record other than those that diplomatics and archival science have considered so far?*

There was considerable difficulty in responding to this question adequately, mainly due to the fact that the objects or entities studied in the diplomatic analyses are not always the same as those that are identified as the digital entities under study in the final reports or other documentation on the case studies. Another factor that muddied the individual responses of the case studies to this question is the fact that in a good number of case studies, the creator believes that his or her documents are records, despite the fact that these documents do not meet all of the criteria making up the definition of a record as established by diplomatics and adopted by InterPARES 1.

Given that the definition of a record adopted by InterPARES 1 relies on the defining criteria of traditional diplomatic analysis, it was the diplomatic analyses of the case studies that prevailed and were used as the definitive source material to respond to this question. In regards to the perception by the creators as to whether their documents were records, despite the archival principle that "whatever the creator treats as a record in the course of any given action is indeed a record in the context of this action," the diplomatic analysis took precedence in determining if the digital entities created in each case study were indeed records. "A record is whatever the creator treats as a record, but that 'whatever' must be something that the creator can in fact keep, associate with other records and subsequently recall."[82]

Artistic sector (Focus 1)

In the arts focus, many of the creators were either unfamiliar or unconcerned with archival terms, notably the definition of "record." As an extreme example, performance artist Stelarc considers that the primary record of his work is his own body. However, most artistic creators are more exemplified by the team that put together *The Danube Exodus*. In this case study, none of the subjects were familiar with archival terminology, but all seemed to make the distinction between works, files used in the actual installation and supporting documents or documents created as a by-product of work production and seemed to be willing to see the latter as records.

In other words, there was some vague sense of a difference between the end product (the work) and the by-products (records), but this distinction and realization was provoked by the questions and intervention of the InterPARES researchers and was not one that the creators were in the habit of making in the course of their activities. After discussing their documents with an InterPARES researcher, the creators were able to see how the end products or works would not be considered records, while the by-products of the action of creating the artwork or installation should be.

[82] Duranti and Thibodeau, "The Concept of Record," op. cit., 32.

The entities studied in three of the ten arts focus case studies fulfil the InterPARES 1 definition of records. The documents kept on Stelarc's Web site satisfy the requirements of a record. The digital entities studied are by-products of Stelarc's performance activities and are bound to each other during the creation of projects, and all records stored on the Web site are bound to each other. For WGBH, both the original footage and the original footage logs posses all the elements required of a record and therefore are both considered records. WGBH must preserve the original footage, the original footage logs and the links between the two entities. In the case of *Obsessed Again...*, the computer code (the MSP/Max patch) is a record because it is the by-product of the act of musical composition, is fixed on the composer's hard drive and is set aside by the composer for future use or reference. The diplomatic analysis did not examine any of the other digital entities created during the composition of the piece.

In the rest of the case studies in this focus, the entities studied were not records or by-products, but rather publications or end products. In these case studies, there is a focus on the final outputs of the activity rather than on the so-called supporting documents. Publications may use or include records, but they are not records themselves. In the arts, these non-records are mainly Web art productions belonging to the category of artwork creation. Although these documents are considered to be records by their creators and meet most of the requirements of the definition, the diplomatic analyses state that they are not records because they are not by-products but clearly end products of creative activities. Therefore, they have been defined as publications.

A good example of this is the case of Arbo. The diplomatic analysis concludes that the group's *Ludosynthèse* is an autonomous entity and not generated as the by-product of the group's research or performance activities. Instead, original records are being modified and edited for publication in the *Ludosynthèse*. As such, the *Ludosynthèse* is therefore not a by-product, but an end product. The Web site has been specifically constructed as a publication, rather than having been set aside during the course of any of the activities documented. Other examples are Altair4, in which the DVD produced is clearly an end product, and *The Danube Exodus*, for which the diplomatic analysis reveals that the exhibit itself and its interactive database are not records because they are end products.

Despite the fact that the digital entities studied in only three of the ten case studies in the arts focus can be considered records, many of the "failed" case studies nonetheless create partial or potential records. The three case studies from the film industry, whose documents were not considered records, provide good examples of this. With Altair4, although the actual film is a final product or publication and therefore not a record, all of the digital entities that are created as the by-products of its creation are set aside and so possess an archival bond with one another and a stable form and content. The National Film Board of Canada presents a near exact scenario. Although each film is considered a final product/publication, all of the digital entities that are created as the by-products of its creation are set aside and thus possess an archival bond with one another and a stable form and content. Comstudio is in a similar situation to the previous two creators, although it is much more focused on the final product, to the detriment of the by-products. In fact, interim documents are "not important" to this creator and are only seen as stepping stones to the final product. Most are not even kept. The by-products of production are considered ephemeral and "exist solely to advance the work of the production."[83]

[83] James Turner et al. (2004), "InterPARES 2 Project - Case Study 09(3) Final Report: Digital Moving Images - Commercial Film Studio," 9. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_final_report.pdf.

Another factor that results in presence of potential records among the creators' documents is the fact that the diplomatic analyses do not always analyze or consider all of the documents created in the given activity, particularly those considered to be "supporting documents." For example, in the case of *The Danube Exodus*, records are created in the process of creating the exhibit (such as Forgac's notes), installing the exhibit and in the collaboration among the various contributors. The diplomatic analysis does not assess whether the "supporting documents" are records, most likely because there is too little information in the case study interim report about these types of documents to be able to make an informed decision. *Waking Dream* also creates potential records during its activity. These include the original video footage; the edited video that is projected onto the screen during performance; the original sound samples; the different versions of the edited soundtrack; the computer code used to operate the remote control dowser; the PowerPoint file used to switch between videos; and administrative records related to the funding, planning and promotion of the work. Lastly, although Electronic Café International's documents may not currently satisfy all requirements of a record as defined by InterPARES 1, the diplomatic analysis for this uncompleted case study has concluded that the artworks created in telecollaborative works are themselves records.

In the case studies of the arts focus that may potentially contain records, there is often the lack of one particular, necessary record: a "script" or enabling record to provide the link among the other entities for re-creation and to provide a context for them or instructions for their interaction. It is often the individual responsible for the creative vision who is the sole person to know how all the pieces fit together to interact and work, without there being any documented means of transmitting this knowledge. In large organizations, these instructions may be documented, but individuals often do not feel the need to do this and the resulting lack of documentation may pose a problem.

With regards to *The Danube Exodus*, it is difficult to ascertain whether it will be possible to re-create the exhibit in the future. In short, it appears that it may be a challenge due to the choices made in short-term preservation strategies. Each contributor has varying financial resources and interest in terms of archival preservation. Although *The Danube Exodus* is a collaborative art exhibit, from an archival perspective there appears to be no collaborative effort on how to preserve the work as a cohesive whole for future re-presentation and no existing set of instructions that would enable anyone other than the three creators to reproduce the work. For *Waking Dream*, the creation of an additional record is required if the piece is to be re-performed. Currently, no description of the piece exists, except for a brief description of the general idea behind the performance that is posted on the Web site. For *Waking Dream* to be re-created by other performers, a document will have to be created that describes the characteristics of each of the components and explains how they work together.

Scientific sector (Focus 2)

Whereas the creators in the arts focus had no idea or functional concept of "record," the creators in the science focus had a different definition or concept of "record" than that used in archival science. In the MOST case study, the researchers use the word "record" as a field value in certain contexts. Yet, in other contexts, its meaning and use is more closely aligned with the archival concept of the term. In the case of the Archaeological GIS study, the creator believes that the GIS is a record only after the findings and data in it are published in a journal article or monograph. To the scientists in the NASA case study, records—or "data products" as they are often called—are characterized primarily by the type of data they contain and the degree to which the data are processed; hence, they distinguish between *experimental data records* (EDRs) (i.e., "the measurements received from the instruments at the mission ground control system"), *reduced data records* (RDRs) (i.e., "processed EDRs") and *engineering data records* (EDRs) (i.e., entities consisting "of raw data in the form of packets containing time-ordered sequences of science data obtained by a given instrument together with engineering information that allows instrument teams to check operations of its instruments").[84]

In addition to distinguishing records base on their content and level of processing, the scientists in the MOST case study further characterize records by the ease with which they can be recreated; thus, while raw data files, which are considered the most important files, are never deleted, the scientists state that "any other files may be deleted because it is possible to recreate them." This seems to imply the raw data files are the only digital entities that the scientists in this study consider to be records or, perhaps more precisely, the only records worthy of permanent preservation.[85]

Yet another distinct concept of record is found in archaeology, where archaeologists commonly refer to the "archaeological record," which, in the words of the creator in the Archaeological GIS case study, is understood to mean "the stuff we have out there in the ground."[86] In this sense, the archaeological record is conceived as a theoretical compilation, rather than as tangible, discrete units of information. On the other hand, the Society for American Archaeology (SAA), in its *Principles of Archaeological Ethics*, refers to the "archaeological record" more broadly as including "in situ archaeological material and sites, archaeological collections, *records* and reports."[87] Meanwhile, later in the *Ethics*, archaeologists are encouraged to

> ...work actively for the preservation of, and long term access to, archaeological collections, *records*, and reports. To this end, they should encourage colleagues, students, and others to make responsible use of collections, *records*, and reports in their research as one means of preserving the in situ *archaeological record*, and of increasing the care and attention given to that portion of the *archaeological record* which has been removed and incorporated into archaeological collections, *records*, and reports.[88]

This double use and meaning of the word "record" permeates the archaeological literature and professional discourse and, to a degree, appears to have been a source of some confusion for the creator in the Archaeological GIS case study, as is suggested by the extended discourse on the definition of the concepts of record and fixity, in an archival sense, spoken to the creator by one of the InterPARES researchers during the case study interview.[89]

Finally, another factor that appears to influence the way that many scientists relate to the concept of record, especially in relation to the notions of fixed form and stable content—two

---

[84] In fact, the National Research Council Committee on Data Management and Computation (CODMAC) defines eight distinct processing levels or types for Space and Earth Science Data Records: raw data, edited data, calibrated data, resampled data, derived data, ancillary data, corrective data and user description data (National Research Council, Committee on Data Management and Computation, *Issues and Recommendations Associated with Distributed Computation and Data Management Systems for the Space Sciences*, report no. NAS 1.26183026 (Washington, D.C.: National Academy Press, 1986), 31–32. Available at http://hdl.handle.net/2060/19880017724.

[85] Ballaux, "Case Study 26 Final Report," op. cit., 13.

[86] InterPARES 2 Project - Case Study 14: Interview B Transcription, lines 442–443 (unpublished).

[87] Keith W. Kintigh (1996), "Principle No. 1: Stewardship," in *SAA Principles of Archaeological Ethics*. Available at http://www.saa.org/Publications/SAAbulletin/14-3/SAA9.html. Emphasis added. See also the similarly broad definitions provided in the Canadian Archaeological Association's *Principles of Ethical Conduct* and in the Archaeological Institute of America's *AIA Code of Professional Standards*. Available at http://www.canadianarchaeology.com/conduct.lasso and http://www.archaeological.org/pdfs/AIA_Code_of_Professional_StandardsA5S.pdf, respectively.

[88] Ibid., "Principle No. 7: Records and Preservation." Emphasis added.

[89] InterPARES 2 Project - Case Study 14: Interview B Transcription, lines 391–409 (unpublished).

characteristics that are fundamental to the concept of record in an archival context—is that, because of the inherently experimental and probatory nature of scientific inquiry, scientists view their activities as resulting in a *provisional* body of work that is continuously subject to revision by themselves and/or by other scientists. Even the raw data, which many scientists consider to be the one product of scientific inquiry most worthy of long-term preservation,[90] are themselves subject to continuous revision as, for example, whenever new and more precise measurement tools or data collection techniques are developed. Thus, even raw data, the most fundamental components of science, are, in effect, themselves subject to obsolescence.

A clear example of this notion of the provisional nature of science data and records is offered by the creator in Archaeological GIS case study, who, when asked about his opinion on the long-term importance of the results of his GIS research, and the records related to that research, expressed doubt that anybody twenty years hence would be interested in or need his data or his results. This belief is based in part on the creator's assumption that GIS technology and research methodology will have changed so significantly within the next twenty years that his data and results will have been rendered completely "obsolete," and in part on his observations of how few of today's researchers actually use or rely on the data and results of archaeologists from 100, 50 or even 20 years ago.[91] According to the creator in this case study,

> There is so much archaeological data that is literally falling off the shelves because the boxes are decomposing that we can't analyze it, that I'm just skeptical that anyone is going to have the time or motivation to try to dredge all this stuff up...I really doubt anyone is ever going to need the raw data.[92]

Despite the wide variability in the way that the concept of record is used and interpreted in the sciences, diplomatic analysis reveals that four out of the five case studies in the science focus are indeed creating records in the course of their activities. Regarding the Archaeological GIS case study, the diplomatic analysis shows that the Coalescent Communities database GIS itself is the authoritative record. The creator treats the various versions of the database as records when they are purposely set aside during the course of business. For the MOST case study, the diplomatic analysis finds that the SDS (science data stream) raw data collected by the satellite and the various data products generated using this raw data are all records. The digital entities studied in the NASA case study are all deemed to be records, as well, and in the case of the National Archives and Records Administration experiment (case study 19), the digital entities comprising the "bill of materials structure," as set aside during the business activities of the originating research partner, as well as the digital entities comprising the test records generated and evaluated during the engineering experiment, all meet the requirements of a record as defined by InterPARES 1.

The one case study that does not create records in the science focus, the Cybercartographic Atlas of Antarctica (CAA), is involved in Web-based scientific activities. The diplomatic analysis established that in spite of the dynamic and interactive environment, every instantiation of the assembled data produced by the entities, or every display generated in response to user inquiries, is autonomous and has therefore been assessed as a publication, as long as it will be hosted and consulted online. Furthermore, the Atlas only partially satisfies the definition of a

---

[90] See, for example, discussion of the results of the general study 09 survey question, which asked participants to briefly describe the elements and/or outputs of their GIS projects that they thought should be preserved for future use or reference and why (Preston, "General Study 09 Final Report," op. cit., 68–71).

[91] Pearse-Moses et al., "Case Study 14 Final Report," op. cit., 28, 36.

[92] InterPARES 2 Project - Case Study 14: Interview B Transcription, lines 1508–1511, 1519 (unpublished).

record because both its content and documentary form are subject to continuous change and because it does not possess an archival bond.

However, the analysis adds that the Atlas has the potential to become a record. When the time comes to set it aside and in so doing to stabilize its content and fix its documentary form, it can become a record. This setting aside fixes the digital entity's documentary form and stabilizes its content[93] and gives the "retired" Atlas an archival bond with other records that are organized in an identifiable documentary context. It also indicates that the Atlas participated in an action and is now filed with other records generated in the same action for further actions or reference. The CAA project also creates traditional records in its administrative activities, but these were not the object of the case study, which focused solely on the Atlas and its production environment.

As with some of the cases in the arts focus, in at least one of the cases in the science focus, there is the lack of an enabling record or set of instructions to provide a link among the digital entities for re-creation or to provide a context for them. The National Archives and Records Administration's experiment encountered a problem in that the STEP (Standard for the Exchange of Product Model Data) file only contains the resultant solid model itself and there was no way to store the construction directions of the solid model records in a neutral format. In other words, there is no "script" describing how to produce the machine parts from the model. This is problematic, because it is unlikely that a new model could be constructed from a preserved drawing that would be equivalent in construction to the original model. To the creators, the construction file is the most important file to preserve, but there is no way to do so in a neutral format.

Governmental sector (Focus 3)

Exactly one half of the case studies in the government focus are creating records. This is the case with the Alsace-Moselle land registry, the Revenue On-Line Service of Ireland and the Singapore Supreme Court, all of which are engaging in traditional, registry-type government activities applied to the digital environment. According to the diplomatic analysis, the ordinances and inscriptions created within the Alsace-Moselle computerized land registry fulfil all the requirements of a record and may be considered as such. Strict procedural and documentary controls ensure that these records are reliable and there are procedural and technological controls in place to ensure the authenticity of the records over time. In the Irish case study, the digital certificates, tax forms and debit instruction forms generated from ROS meet all the requirements of a record. In Singapore, the documents created and set aside in the course of the activities of the Supreme Court in administering bankruptcy proceedings also meet all requirements of a record as defined by InterPARES 1, as do at least some of the documents created by the New York State Department of Motor Vehicles On-line Services System, including the core records (i.e., user profiles) and audit trails.

The remaining four case studies in this focus do not create records, as defined by InterPARES 1. These creators are all in the service activities group. Two have open hypermedia features with fluid and changing data that provide various information, but without a document that is set aside. They therefore lack fixity of content and/or form, which is needed for the entities to be considered records and preserved over time.[94] VanMap cannot be considered a record because: (1) there is no act, (2) it has not been set aside and has thus not acquired an

[93] See Records Creator Principle C1 in Appendix 19.
[94] Ibid.

archival bond and (3) its form is not fixed. Although less fluid and dynamic, the Legacoop of Bologna's Web site contains entities that do not satisfy all the requirements of a record. Specifically, current practices do not ensure stability of content, the entities do not possess an archival bond beyond a chronological record of their posting and the procedural context is under-developed. Furthermore, several documents on the Web site may be considered publications, such as the member's newsletter.

The remaining two cases also create publications, not records. The Antarctic Treaty Searchable Database does not meet the requirements of a record primarily because the database does not participate in an action and does not possess (or require) an archival bond. It is a compilation of documents selected and gathered for dissemination that is developed to stand alone as an information resource and is therefore a publication rather than a record. The representation of the component documents and records in the database does not have the effects and contexts equivalent to the originals. In the Ontario Archives Web exhibits, the files found on the production server do not fulfil certain requirements of the definition of a record. Their autonomous nature reveals them instead to be publications.

Despite the shortcomings of these cases that prevent their documents from being considered records, like the previous two focuses, the "rejected" cases in the government focus nonetheless contain partial or potential records. At the Archives of Ontario, although the files on the production server are not considered records, the files that are stored on the development server fulfil all InterPARES 1 record requirements and may therefore be considered records. They are public records and narrative records of the activity of creating exhibits. In the case of VanMap, the system has the potential to become a record once it has been set aside. This is also the case with the Legacoop of Bologna's Web site. The creator uses the documents on the Web site as records and the site is used as a place to post important documents. These include some documents that are not found elsewhere in the creator's fonds. However, they will not fulfil the InterPARES definition of a record until they are set aside. The act of setting the site aside will stabilize its content and link it to the Legacoop's other administrative records as evidence of its activities.

General observations

As examination of the case studies in each of the three focuses has shown, the diplomatic analyses reveal that digital entities that satisfy the requirements of a record are mainly those created: 1) to support the creation of artwork, 2) to support scientific activities and 3) for use in public filing and registry systems. All of these have open hypermedia features or belong to the second type of digital entities discussed earlier (i.e., digital components like files, program code, etc., that take place in a larger process or action involving other digital and/or analogue means).

Some conclusions can be drawn regarding the documentary form produced. Generally speaking, there are two reasons why documents have not been assessed to be records. In several of the cases, there was a problem related to the capacity to ensure the stability of the content and fixity of the documentary form. In many of these cases, open hypermedia features generated fluid and changing data. In these cases, the creator must ensure there is a fixed store of data within the system and that the rules by which the data are aggregated and presented on screen are predetermined, consistently applied and well documented. Part of the problem stems from the fact that the activities of these creators often give rise to new or emerging productions in which the form may evolve according to the development of the technological approach over time. These cases are strongly linked to the availability of software and the ease of using it in a given field. On that basis, if the creators want to allow the transformation or re-creation of their work, they must keep in mind that their documents will have to remain as independent as possible from the applications used to create them. Under these conditions, they must use software and file formats that offer the best hope for ensuring accessibility of the records over time. Software that is not compatible with previous versions (backward compatibility) or with future versions (forward compatibility) impedes accessibility over time. Creators must also ensure that software for one application works well with that of other applicable applications and systems (interoperability).

In many other cases, information technology was used to create end products dedicated to the dissemination of artistic works or scientific knowledge. These are self-contained entities that stand on their own and do not require any other information to be understood. This situation leads the creators to a false sense of completion. Arbo is a good example of this. Their Web site integrates documents judged to be the most representative of fifteen years of performance of their theatre troupe. The artists refer to it as a tombstone, since for them the site has become a repository of their entire memory. They therefore did not feel the need to link it to the rest of their fonds. However, the fact is that although Web art is open work, it often has a predetermined finishing point in time. The work may remain open for a fairly long period of time, but it is eventually closed. In the same way, both scientific and non-administrative governmental end products are eventually retired. At that point, the documents should not only be stable, but they should also have archival bonds with other traditional and digital records in the creator's fonds. This implies, therefore, that at a minimum, creators must be aware that digital records should be organized into logical groupings consistent with the organization of the paper files and linked to retention periods as much as possible.

Table 6. Digital Entities' Fulfillment of the Criteria of a Record

CS#	Entity Studied	Criteria Necessary to be Considered a Record				
		Fixed Content and Form	Participate in an Action	Archival Bond	Three Persons	Identifiable Context
1	*Ludosynthèse*	√	X	X	√	√
2	Documents on Web site	√	√	√	√	√
3	Online issues and database	√	X	X	√	X
	Component documents[95]	√	√	√	√	√
5	Files on production server	√	X	X	X	X
	Files on development server	√	√	√	√	√
6	Cybercartographic Atlas	X	X	X	√	√
8	Data (raw data and SPICE files)	√	√	√	√	√
9(1)	*House of Julius Polybius* DVD	√	X	X	√	X
	Individual production files	√	√	√	√	√
9(2)	Animated films	√	X	X	√	X
	Component documents[96]	√	√	√	√	√

[95] Computer programming code, graphic design, artist commissions, editorial scripts, database architecture, etc.
[96] Including textual, artistic, database and musical components.

9(3)	Digital animated films	√	X	X	√	X
	Digital moving image material	√	√	√	√	√
9(4)	Original footage and footage logs	√	√	√	√	√
10	Component documents[97]	√	√	X	√	X
12	Antarctic Treaty Database	X	X	X	√	?
13	Computer code (patch)	√	√	√	√	√
14	Coalescent Communities DB	√	√	√	√	√
15	*Waking Dream* performance	X	X	X	√	√
	Public Web site	√	X	X	√	√
	Component elements.[98]	√	√	√	√	√
17	User profiles (core records)	√	√	√	√	√
	Audit trails	√	√	√	√	√
	Licenses/registrations	√	√	√	√	√
	Driving record abstracts	√	√	√	√	√
18	Ordinances and inscriptions	√	√	√	√	√
	Entire database	X	√	X	√	√
19	"Bill of materials structure"[99]	√	√	√	√	√
	Test records (NARA experiment)	√	√	√	√	√
20	Digital certificates	√	√	√	√	√
	Tax forms	√	√	√	√	√
	Debit instruction forms	√	√	√	√	√
21	Bankruptcy records in EFS	√	√	√	√	√
24	VanMap GIS system	X	X	X	√	√
25	Documents on Web site	X	√	X	√	X
26	SDS raw data and data products	√	√	√	√	√

Question 5

As government and businesses deliver services electronically and enter into transactions based on more dynamic Web-based presentations and exchanges of information, are they neglecting to capture adequate documentary evidence of the occurrence of these transactions?

This response revealed a cleavage between the arts and science focuses on one side and the government focus on the other side. The first two fields capture little or no documentary evidence, due to the fact that the creators do not view themselves as participating in

[97] Forgacs' notes, video and multimedia documents, still images, texts, etc.
[98] Including sound samples and video recordings, compiled soundtrack, computer code and PowerPoint file.
[99] From the originating partner.

"transactions," which was largely understood as the provision of goods or services. They simply did not feel that the question applied to them or that they had the legal obligation to capture documentary evidence. On the opposite end of the spectrum, with few exceptions the creators in the government focus have sophisticated means and procedures for capturing documentary evidence, which is seen as ensuring the evidential value of the documents that they produce.

Artistic sector (Focus 1)

This question does not apply to the majority of case studies in the arts focus. This is mostly because the case studies in this focus do not involve the delivery of services (unless a performance may be considered a "service"), or the creator does not enter into transactions. However, another fact distinguishes this focus from the other two. For individuals and small businesses in the arts focus, there is no legal mandate to keep their records, but rather only an interest on the part of the creator to sufficiently document his or her activities to be able to re-create or reproduce the work.

HorizonZero states that for most of the activities undertaken by the team, there is no legal mandate to make or keep documentary evidence, a sentiment that is echoed by performance artist Stelarc, the composer of *Obsessed Again...* and the *Waking Dream* team. However, it is clear that if the creators of *Waking Dream* intend that the work continue to be performed in the future—and in particular, if they hope to have performers other than themselves do so—better documentation of the team's intentions and of the methods and technologies used to realize those intentions will be required. It is also in the best interest of the composer of *Obsessed Again...* to document his process of creation and the characteristics of each element of his work as completely and accurately as possible to facilitate future performances of his work, especially if the accurate reproduction of his intentions is important to him. Despite the lack of a legal mandate to capture documentary evidence for most artistic creators, there is still the question of whether adequate documentation is being kept in an organized and systematic manner to allow for the recreation of the installation of the work or performance as a whole.

For larger businesses or creators in the arts focus, there is still a lack of documentation of records during the creation process. What documentation is being done in this creator context may depend not on institutional or official policies, but rather on the practices of individual artists or producers. At the National Film Board of Canada, little, if anything, is documented about the creation of the film itself. The way in which entities are documented and maintained depends on the individual supervisor for each production and the size of the project (small projects may only be sparsely documented). In fact, problems have been encountered in post-production due to insufficient or poor documentation. In the case of Comstudio, new artwork is not so much created as it is merged with additional artwork to form a single file. Previous versions are saved for a time in case it is necessary to consult previous iterations, but eventually old versions are overwritten and no documentation of their existence or transformation is captured.

The individual practices followed for documentation in the arts focus mostly include naming conventions. At the National Film Board of Canada, every production is assigned a number and this same number applies to all the documentation concerning the production, both paper and digital records. At Comstudio, naming conventions are used to identify digital entities and provide information about sequence, scene, name of the object and numerical information to identify the version. Physical images are numbered and once they are scanned into the Avid computer, these numbers appear and are used to link the digital images with the images on paper. At WGBH, a unique identifier links the catalogue record in the footage log with the original footage. Both the original footage and the original footage logs follow naming conventions that allow them to be linked together and to the final program production. It should be noted that these three examples, which all come from the motion picture industry, are among the most organized and institutional of the creators in the arts focus. Smaller organizations and individual artists may not even apply these basic documentation measures. "For individuals, the burden may seem great, but the alternative—loss of records or the emergence of corrupt and unverifiable data—would be an even greater problem in the long run."[100]

Scientific sector (Focus 2)

As with the arts focus, this question does not apply to (or there is "no legal obligation" for) four out of the five case studies in the science focus. Once again, the primary reason is that no transactions occur in the activities of the creators studied. For example, although it is an agency of the United States government, the focus of NASA's work is scientific endeavour; as such, no transactions occur. The transactions for the users of the Cybercartographic Atlas of Antarctica are primarily to view and interact with the content for educational purposes. In this context, there is no legal obligation to maintain a record of these transactions. Nonetheless, from an historical perspective, there is an interest on the part of the creator in preserving the entire CAA at different points in its development.

The only case study where this question seems to apply is that of the National Archives and Records Administration of the United States. In this case study, a product data management system captures all of the digital entities within the scope of creating a digital solid model (the entities that are created in the CAD system). The product data management system captures all actions and transactions that take place within the system. There is a rigorous change-control process, whose changes are recorded in the product data management system. In this case, the term "transaction" may be understood to be an action, modification or procedure performed on the digital entities and not the exchange of money, information or goods.

Governmental sector (Focus 3)

The government focus is the opposite of the other two focuses in regards to this question. In fact, the question applies to all but one case study in the focus (The Antarctic Treaty Database). Most of the case studies in this focus have extensive login capabilities to record or capture all actions and transactions performed in the system. These cases are mostly implementing electronic versions of traditional registry or service activities. The evidential weight of the records that they create depends upon capturing this information and assuring the accuracy, authenticity and reliability of the system's records. These requirements are reflected in the relevant legislation governing the activities in question.

For example, the Alsace-Moselle land registry has extensive login capabilities for recording all actions and transactions taking place in the system. It is extremely thorough in terms of capturing documentary evidence. Transactions are conducted within the system itself and information relative to all properties is contained within a database. In Ireland's Revenue On-Line Service, transactions are documented through the formal act of signing and submitting a tax form to the Revenue Commissioners via the ROS. This is considered evidence of a record. In Singapore, the Bankruptcy Act and Bankruptcy Rules ensure the continuation of strict procedural controls over all transactions conducted by the government, even in the electronic realm. Transactions are also strictly controlled by the rules of the court, internal work processes and

[100] See the *Creator Guidelines* in Appendix 20.

Practice Directions. In fact, the electronic system is capturing more documentary evidence than the paper-based system. This is likely the case also with the New York State DMV's On-line Services System, which maintains extensive user logs and audit trails that track all transactions and changes made by both employees and customers so that records cannot be modified without leaving behind evidence of that modification. The DMV also makes use of a strict access rights system that controls what type of access each employee has to the digital entities in the system. As well, all online DMV transactions are electronically transmitted using Secure Socket Encrypted Transactions that are authenticated through the use of digital watermark technology.

Some exceptions to this trend of thorough government documentation involve case studies in which changes or transactions are undocumented or in which data are overwritten. In the case of the Ontario Web exhibits, changes to Web exhibits may be made by the creator without consistent, or even any, documentation. There was also no indication of any documentation around reformatting older exhibits to bring them to current standards. Interviewee comments confirm that recordkeeping of supporting documentation is done individually in terms of what is created and captured, and how and where it is filed. In the case of VanMap, different data are updated at different times, either on a regular basis or as needed. Most updates consist of data being overwritten. For data that are overwritten, there is no way to track updates over time. There is also no capture of actions or transactions in the system. It should be noted, however, that these two examples are non-traditional government activities and resemble case studies in the arts and science focuses, respectively.

There are also two exceptions in which documentation is less thorough than normal for government creators, due to the fact that the creator's electronic system is not connected to the traditional, paper-based system. For the Ontario Web exhibits, these exhibits are not treated the same as traditional, physical exhibits. The recordkeeping process described in one institution in the Exhibit Approval Form has evidently not been followed. In the case of the Legacoop of Bologna's Web site, there is no doubt that the dynamic, Web-based entities are neglecting to capture adequate documentary evidence. This may be due to the fact that the creator does not focus on the Web environment with the same quality and attention as is given to traditional documents. Although a recordkeeping system is used, digital records are not considered to be part of any formal recordkeeping system. Some documentary evidence, therefore, is being neglected.

Question 6

> *Is the move to more dynamic and open-ended exchanges of information blurring the responsibilities and altering the legal liabilities of the participants in electronic transactions?*

As with the previous question, this question was seen to not apply in large part to creators in the arts and science focus groups. Legal liabilities in the arts focus are largely limited to intellectual property considerations, which may become complicated in a partnership or multiple-creator situation. In the science focus, professional ethics and norms seem to define the bulk of creators' responsibilities and liabilities, while the legal liabilities of government creators delivering traditional registry-type activities in the digital environment have largely remained unchanged by the move to a technological environment.

Artistic sector (Focus 1)

This question does not apply to many of the case studies in the arts focus. In some cases, such as Altair4, the question is not applicable because the case study does not involve electronic transactions. This is the situation for several case studies in this focus. For others, the lack of relevance for this question hinged more on the fact that the artist does not have any legal liabilities, as expressed by the composer in *Obsessed Again...*, or on the perception that the artist, such as performance artist Stelarc, does not have to live up to legal requirements in the same way that government or businesses have to.

When artistic creators do recognize legal requirements or responsibilities, these are all copyright or privacy related. Some creators must obtain a rights release prior to reproducing certain works, while for others a contract states that individual artists retain the copyright to the work, but the publishing or performing body retains the financial or re-use rights.

The attention to and understanding of these rights and responsibilities varies widely in the field, based largely on the level of complexity and sophistication of the creator. Arbo knows that it is subject to copyright and disposition laws, but the legal questions concerning photographs are only vaguely understood and addressed by the group. In *The Danube Exodus* project, these issues are aggravated as a result of the complicated authorship and ownership of the work and its component parts, including those to which institutions other than the primary authors hold copyright. *HorizonZero* states that there are no specific laws or regulations governing its activities except for those related to copyright, which are specified in the artists' contracts. At the National Film Board of Canada, complexities exist with respect to film rights due to the contracting out of work, while at the uncompleted case study, Electronic Café International, ECI owns the rights to its telecollaborative works, although certain performances cannot be recreated until third parties have released their rights.

Despite the sometimes complex and misunderstood issues of copyright and privacy in the arts focus, these responsibilities appear to be consistent with a paper-based environment and remain unchanged with the move to a digital environment. At *HorizonZero*, the case study final report does not provide any information indicating that copyright issues have been affected or altered through the use of digital technologies. A similar situation is encountered in most of the other case studies, such as *The Danube Exodus*, where there is nothing in the case study documentation that specifically addresses how this situation would differ from a similar situation in a non-digital environment. Even among the *Waking Dream* team, where there is conflict over whether the work is performance art (therefore proprietary) or theatre (which would be reproducible), "this issue revolves around the nature of the performance, not the nature of the technology used to create it and would manifest itself in the same way in a non-digital environment."[101]

The greater the hierarchy involved, the more formalized the rights are in the arts focus. Loose associations of individuals seem to recognize and respect each other's rights, without these being formally spelled out or documented. Most responsibilities or obligations for individuals and small groups come from moral or ethical concerns or from granting or funding bodies, rather than from legal issues. For example, at Arbo, the group's ethical code requires that the photographer's name is identified for each work and the group has agreed to never force a spectator to participate in a performance. Also, the *Waking Dream* team claims to have no legal liabilities or responsibilities besides compliance with grant stipulations.

[101] Douglas, "Case Study 15 Domain 1 Research Questions," op. cit., 4.

Small businesses and agencies formalize their legal rights and responsibilities more by entering into contracts of varying complexity. The uncompleted case study, Electronic Café International, is bound by the contracts into which it enters with contributing artists. Similarly, *HorizonZero* negotiates a copyright license with each of its contributors, whereby the contributor retains copyright over the work but waives financial rights, so that the Banff Centre has rights in perpetuity for the reproduction of the work in digital or print format.

Larger groups and businesses have given more thought and resources to legal responsibilities and digital security. This is most evident in the motion picture industry. At the National Film Board of Canada, copyright law applying to the use of third party segments in films led to the development of an electronic rights management database, which is now an integral part of the creator's Synchrone system. WGBH also uses a digital asset management system, which includes a login procedure so that archives personnel can track use of the digital library and secure certain assets. Any changes are tracked by system administrators and only archives personnel can modify metadata information linked to the footage. Comstudio has put in place strict internal security controls to ensure that responsibilities and rights are clear. An approval process exists to provide access to specific files. In short, while individuals and small companies are "putting out fires" (i.e., dealing with problems as they arise), larger companies and government are installing firewalls (i.e., taking pro-active actions to prevent problems). For example, Arbo has agreed to remove any image from the *Ludosynthèse* as a result of complaints received from those who appear in them, while the large motion picture studios are organizing their assets into a digital rights management system to prevent any future problems from arising.

Scientific sector (Focus 2)

In the science focus, most of the creators' legal responsibilities or obligations come from legislation that is already in place for a paper-based environment, as well as from professional standards and obligations. For example, NASA is an agency of the United States government and, as such, must adhere to the governing legislation from which it draws its mandate. In the case of the Center for Desert Archaeology, major federal regulations and policies affecting the protection and management of archaeological resources in the study area are embodied in several laws, regulations and executive orders, which impact on the group's activities.

As in the arts, the realization of, and response to, legal responsibilities increases with the size of the creator in the science focus case studies. Due to its size, structure and organizational culture, the Center for Desert Archaeology does not rely on procedures in a formalized sense—instead, they are inferred. Even in the MOST project, there are hardly any procedures (with the exception of the MOST Archiving Manual), due to the organizational culture, the size of the research team and the resources available. On the Cybercartographic Atlas of Antarctica Web site, the group's responsibilities are disavowed with disclaimers as to the accuracy and reliability of information for other than educational purposes.

In some of the science focus case studies, issues of professional ethics come into play in the question of rights and responsibilities. At NASA, those involved in the creation of documents are bound by professional ethics in the planetary sciences as well as the institutional ethics of NASA, which call for trustworthiness and competence. Even in a small creator context such as the Center for Desert Archaeology, there are no overwhelming ethical issues that arise in this research activity on a daily basis, but there are many overriding professional ethical concerns that govern certain practices within the North American archaeological community.

For some case studies in the scientific field, memoranda of understanding seem to replace the contracts that were noted in the arts focus. These documents outline the responsibilities and obligations of each party. For example, in the National Archives and Records Administration experiment, the three trusted research partners are bound by memoranda of understanding, in which responsibilities are outlined in terms of the engineering/archival experiment.

Governmental sector (Focus 3)

For the most part, legal liabilities of transactions in the government focus have not been altered in the move from a traditional to a digital environment. In Alsace-Moselle, the judge is still personally responsible for the verification process and may be sued by the state if errors are made. The judge has sole competence for the creation and signature of ordinances and thus for inscriptions within the registry, even the computerized version. In Ireland, Revenue is still responsible for the collection and management of taxes. It must still adhere to the legal mandates related to Irish law and to Ireland's membership in the European Union. Legal liabilities of transactions have not changed in the case of the Singapore Supreme Court, either. In the digital as in the traditional environment, the court maintains its role of information service provider for establishing the creditworthiness of individuals and for setting legal precedents. The same also holds true for the New York State Department of Motor Vehicles, who is still responsible for issuing, renewing and replacing vehicle licenses, registrations and titles.

For formal, registry-type activities such as the ones exemplified above, there is a more heightened recognition among creators in the government focus of the obligation to create a secure Web environment needed to support transactions. There is often strict control over access, such as through the use of PKI and biometrics, to protect privacy and confidentiality. In Alsace-Moselle, the judges, in particular, had deep security concerns in case of tampering or system malfunction, due to their heightened responsibilities. In Ireland, Revenue currently remains aware of its responsibilities and legal liabilities regarding the digital records generated in the Revenue On-Line Service, which must comply with the E-Commerce Act of 2000, while the Supreme Court of Singapore ensures strict control over access to, and use of, its records to protect the privacy and confidentiality of involved parties. Likewise, the On-line Services System of New York State's Department of Motor Vehicles is subject to numerous state and federal laws and regulations dealing with issues such as system security, protection of privacy, the use of electronic signatures and requirements regarding accessibility of online DMV services for persons with disabilities.

Smaller groups, or those engaging in less traditional activities, often disengage their responsibilities with disclaimers as to the accuracy and reliability of information for other than educational purposes, similar to what was seen with the Cybercartographic Atlas of Antarctica in the science focus. It should be noted, however, that these case studies are more like those in the science focus than their peers in the government focus. At VanMap, whenever the public version is opened, a disclaimer appears reading, in part, that "The City assumes no obligation or liability for the use of VanMap by any person and makes no representations or promises regarding the completeness or accuracy of VanMap or its fitness for a particular purpose."[102] The user is required to click OK to use VanMap. For the Antarctic Treaty Searchable Database, The end-user license agreement included with the webCDserver versions of the database includes a disclaimer against the accuracy and reliability of all documents included in the database. Any inaccuracy in the copies included in the database is attributed to the creators of the original documents.

[102] McLellan, "Case Study 24 Final Report," op. cit., 17.

Some of the governmental creators engaging in non-traditional activities have fewer or no defined standards or responsibilities. In the case of the Ontario Web exhibits, there are no internal policies at the Archives of Ontario governing the creation, storage, or access to Web exhibits. The *Management of Recorded Information Directive* (the existing Ontario Government recordkeeping policy) is not adhered to (e.g., Web exhibits are not governed by a records retention schedule). Similarly, the policies, procedures and standards used to determine how to include and present data in VanMap are not extensively documented. At the Legacoop of Bologna, the responsibilities are not clearly identified with reference to control of the integrity of digital entities. The recordkeeping system in place has no relationship with the cooperative's Web site. What is more, Legacoop is not bound by any formal obligations, short of an ethical obligation to ensure that information posted on the Web site is accurate and correct.

Question 7

> *How do record creators traditionally determine the retention of their records and implement this determination in the context of each activity? How do record retention decisions and practices differ for individual and institutional creators? How has the use of digital technology affected their decisions and practices?*

When considering the response to this question before looking at the evidence from the case studies, one might presume that the experience of using digital technology would lead creators to greater appreciation and understanding of the need to make retention decisions and implement preservation procedures based on the technology that they are using. Whereas paper records might last indefinitely in a traditional environment despite a lack of formal retention and preservation policies, an informed awareness of the tendency for digital documents to become inaccessible due to technological obsolescence might compel creators to be more proactive in a digital environment and consider the retention needs of their organization to preserve their records. At this point, the discussion now turns to the activities and practices of the case studies in the three focus groups to see if reality is indeed in line with this presumption.

Artistic sector (Focus 1)

In five of the case studies in the arts focus, the creators have either not considered record retention adequately or at all, or have no formalized criteria if they have considered the question. Part of the problem stems from a lack understanding of archival practices, of course, but another factor is the lack of clear definition of responsibilities in a partnership or group setting, such as in *Waking Dream*, whose creators have not considered record retention, or in collaborative efforts such as *HorizonZero*, where no formal recordkeeping procedures have been identified.

In some of the case studies that have not considered records retention or that do not have formal retention policies, the lack of a policy, ironically, does not lead to a lack of retention. In fact, in such cases, since the creator has not determined what to keep, they instead keep everything, which ensures that they keep the important records (along with everything else, of course). In the case of WGBH, the final report does not give information on specific retention times, but does indicate that the creator maintains footage dating back to the 1950s. Perhaps it can be assumed that the creator retains all of its production work related to the television programs it produces. At Altair4, selection is made on the basis of importance and similarity. For example, when two versions are practically identical, only one is saved. This creator therefore saves about 90% of the digital entities that it creates. As for Arbo, the absence of criteria results in no real established standard in place for preservation. Since the group's members have no defined selection criteria, they therefore keep everything.

For the artistic creators that do have criteria or policies to guide their record selection and retention, the majority of them keep records for their own business needs (which vary from creator to creator), or for future re-use or re-purposing. In other words, since there is no legal mandate for most creators in the arts focus to create records, it may be assumed that a record's retention period or the decision to preserve records is determined based on the needs of the creator for access to his or her own records for use and/or reference, if such a decision is indeed actively made. However, selection criteria are highly subjective and may be unspecified or undocumented.

For example, the composer of *Obsessed Again...* has not considered the issue of record retention as such, although he is concerned with the ability to perform his work in the future. Performance artist Stelarc chooses records for their convenience and whether he thinks they are effective publicity tools for his continuing activities. At Altair4, retention decisions are made based on legal and marketing reasons. Files are retained when there is a specific reason to do so, but there is no defined retention schedule in place. Some creators, such as *The Danube Exodus*, have priorities for preserving certain documents and digital entities over others. The works themselves—or parts of those works—are preserved, along with the files needed to render them. Next in importance are documents that describe or illustrate how the installation should "look, work, behave."[103] Least important in terms of long-term preservation are administrative records, including meeting minutes and correspondence. At Arbo, some activities require the preservation of documents to reintegrate them into subsequent performances. Again, subjectivity was the most important factor in guiding the selection of records for the creation of the *Ludosynthèse*.

In some cases in the arts focus, the selection of documents for preservation was driven—even limited—by the technology used in the creation process. The creators at Arbo realized that their digital entities could only be read using the specific programs with which they were created. These programs would therefore need to be maintained to access or use the entities in the future. For Stelarc, the selection of records to be posted to the Web site and thus retained is often technologically driven, as well. *HorizonZero* attempted to get around the problem of obsolescence by transferring all files related to the project to a single personal computer capable of running all of the software and hardware needed to access all file types used in the project.

Scientific sector (Focus 2)

As with the arts focus, creators in the science focus generally keep their records for future business needs or usability. At NASA, the Planetary Data System was designed for long-term preservation and usability of data. A similar situation can be observed in the National Archives and Records Administration experiment. The business owner that created the original documents must be able to access and use his or her records for business purposes over a long period of time (over fifty years).

Conversely, in the science focus documents and digital entities are often discarded once they are no longer useful. In the case of the Center for Desert Archaeology, there are many intermediary files that are created during the course of the research that are discarded once the calculation is completed or the research question has been fully answered. At the MOST project, only the FITS and SDS files are routinely captured and backed up. From the moment that other entities are not up-to-date (because there is a better reduction, for instance) they can be removed

[103] Hubbard, "Case Study 10 Final Report," op. cit., 6.

from the system. An exception to this practice is the Cybercartographic Atlas of Antarctica, which uses an open source content versioning system to capture, track and backup all versions of its code.

The retention and organization of entities in the science focus often reflects the creation or business process. At the Center for Desert Archaeology, the creator retains the digital entities, but not in a separate, formal recordkeeping system. Rather, the ad hoc documentary procedures mirror the business procedures. The digital entities are usually organized by project, which mirrors the majority of the creation process, which itself is project- or problem-focused. Data are usually created in relation to a project and the filing schema is not significantly altered once the project has been completed. This situation is also seen at MOST, where the digital entities are organized by target (star) and date and thus reflect the creation process. VanMap also follows this procedure. In this case, the organization and schema are dictated by the nature of the activities used to create the data.

One surprising fact concerning records retention by creators in the science focus is that Microsoft Windows tools are often used as the only recordkeeping system or means of capturing the digital entities that are retained. In a field where advanced and cutting-edge technology is used to measure variations in the brightness of stars thousands of light years away and to map out and present dynamic information from Vancouver to Antarctica to the Arizona desert, similar trend-setting technology is not being used to capture and maintain the results of these scientific activities. This fact holds true in creators of all sizes, from small, private groups to large, government projects. At the Center for Desert Archaeology, other than elements of the Microsoft Office Suite, there are no collective capture tools for the information within the GIS. Similarly, at the MOST project, other than Microsoft Windows tools, there is no formal capture system in place. All digital entities are accessed via Windows Explorer. The Cybercartographic Atlas of Antarctica was the exception, using primarily open source solutions and MAC technologies.

To implement decisions made regarding the retention of their records, many creators in the science focus case studies make periodic backups of their data and software. The data are most often backed up to CDs or DVDs. Smaller groups may maintain their data on a team member's personal computer for access, but nonetheless back up their entities. At the Center for Desert Archaeology, besides burning data to CD-ROMs, there are no systematic retention strategies in place. Larger groups may have formal, documented procedures, such as the Archiving Manual in the MOST project. In these larger organizations, all records or data are systematically backed up according to a procedure in place for the creator or the given project. At MOST, the software is also updated and backed-up for obsolescence reasons. Because the FITS files and reductions are done in specific software, the MOST researchers periodically back up the software so that each reduction can be redone in the same software environment in which it was originally created. For the creation of the FITS files and reductions, the old versions of software are preserved.

Migration is also commonly used by the creators in the science focus, but not as an archival tool. Instead, migration is performed as a means of maintaining the usefulness of retained digital entities. In the National Archives and Records Administration experiment, there is some concern that the use of technology to preserve the digital solid model records, such as encapsulating the CAD file into a STEP file, will fail. As a result, a TIFF image of the drawing is also created so that if all else fails, the image of the drawing will survive and the original model can be reconstructed from the TIFF image. In the case of the Center for Desert Archaeology, the organization is actively migrating the files to newer versions of the software. However, this is the most that they are doing in terms of addressing software and hardware obsolescence.

Governmental sector (Focus 3)

In some case studies in the government focus, the move from a traditional to a digital environment has lead to more attention to the issue of digital retention and more formalized procedures or precise rules in this area. The cases in which this is true are those that perform traditional registry or service activities, but through digital means. For the Alsace-Moselle land registry, in the paper-based system, records were retained indefinitely but were not transferred to an archival authority—the land registry offices maintained the registers. Computerization has meant that retention periods must be instituted and records must be transferred to an archival institution. The GILFAM must specify the length of time it will keep the records in the computerized land registry and the method by which it will transfer the records to an archival institution. At the Singapore Supreme Court, one motivation for the implementation of the electronic filing system was to solve the storage problem of paper records. At the same time, the court is conscious that despite the availability of digital storage space, it may be more cost effective and efficient to impose stricter retention guidelines with appropriate checks and balances embedded in the workflow to ensure that documents are deleted as soon as they cease to have value and that only those that require long-term storage are retained in an online or offline environment.

Despite this heightened awareness of the issues involved in the retention of digital records among some of the more traditional-type creators in the government focus, others have either no set rules for digital retention, or the traditional retention rules do not apply to the digital environment. In some cases, the digital entities are not part of the recordkeeping system at all, even if there is the assumption on the part of the public that the creator has a legal requirement to keep records. The most surprising case in which this is true is the Revenue On-Line Service of Ireland. Although Revenue abides by the National Archives Act as a guide to retention practices for all paper-based records and requires authorization to destroy any tax forms, no strategy has yet been articulated to deal with the retention of records found or created within the digital environment. For example, it is unclear for how long older public keys and digital certificates are maintained. In the case of VanMap, the digital entities cannot be said to form part of a recordkeeping system and no preservation strategies are currently being employed in the archival sense. For some of the creators, although there are no preservation strategies in place, there are some individual practices in use, which would almost certainly become components of a preservation strategy were one to be developed and implemented.

Due to the lack of a formalized preservation strategy for many of the creators in the government focus, many digital entities are overwritten or deleted as they are updated or are no longer useful, as was seen in the science focus. This overwriting or deletion is performed without the capture of the previous instantiation of the digital entity. For example, the Revenue On-Line Service of Ireland maintains only a subset of its records, retaining the records only of active users and agents. In the case of VanMap, the geospatial data are generally not captured or kept, but rather are overwritten as needed. The data and the HTML pages are recorded and saved, but are overwritten as needed and previous versions are not captured within a recordkeeping system.

Despite these practices, there are nonetheless some strategies employed by government creators to counter the problem of obsolescence. For example, at the Singapore Supreme Court the outsourcing of digital certificates to a licensed certification authority to counteract technological obsolescence is being considered. However, throughout the government focus, the migration of records is not performed to counteract obsolescence or as an archival tool, but to maintain use, as was also noted in the science focus. In Singapore, migration is recommended only for active and semi-active records (those that have a greater chance of being used), while microfilm is recommended for the permanent storage of court records. In VanMap, not only have the existing data been migrated to the Oracle Spatial database, but the use of the new system is expected to streamline the processes that create the data and to also allow VanMap users to view mostly live data instead of static image files. The purpose of migration was thus not to preserve the entities as archival records, but to improve the use of VanMap.

In the government focus, more than the other two focuses, questions of protecting personal information appear with regards to the retention of records. For the Alsace-Moselle land registry, the French public agency that deals with privacy issues (*Commission Nationale Informatique et Libertés*) mandates that all personal information be destroyed past the period for which it is useful for the purposes for which it was collected, except if its preservation is required for historical, scientific or statistical purposes. In the case of Ireland's Revenue On-Line Service, it has been noted that the preservation of tax records is not appropriate, given the level of personal information within it and its lack of suitability for archival preservation.

PART THREE

AUTHENTICITY, RELIABILITY AND ACCURACY OF DIGITAL RECORDS IN THE ARTISTIC, SCIENTIFIC AND GOVERNMENTAL SECTORS

Domain 2 Task Force Report

by

John Roeder, The University of British Columbia
Philip Eppard, University of Albany, State University of New York
William Underwood, Georgia Tech Research Institute
Tracey P. Lauriault, Carleton University

Introduction

Background and mandate

The first InterPARES Project (1999-2001) addressed the problems of preserving administrative and legal records generated within databases and document management systems.[1] Such records, although fixed digitally on relatively unstable media, are intended to approximate the physical documents generated in the course of established business procedures in well-understood juridical contexts. Thus, the Project naturally focused on how to preserve their authenticity and reliability—those qualities that make them trustworthy as the representations of actions—during their inevitable rewriting from system to system, from medium to medium and from format to format, when they are susceptible to alteration.

For this investigation, InterPARES 1 drew concepts from contemporary archival diplomatics, a theory of record and record analysis rooted in a European practice that is a source of modern Western business and legal systems. Diplomatics identifies those features of documents that make them records—fixed, reliable, complete representations of transactions. It helps guide preservation, because preserving a record requires preserving all those features that make it a record.

Referring to this theory, the Project's Authenticity Task Force developed two sets of practical guidelines for ensuring the authenticity of digital records over time.[2] Each addresses a different phase in the lifecycle of a record, assuming a common distinction between active records that are maintained by the creator for current and future reference, and inactive records that have been transferred to the custody of an archive for long-term preservation. The *benchmark requirements* set forth a basis for presuming or verifying the authenticity of the creator's digital records, while the *baseline requirements* support the production of authentic copies of digital records after they have been transferred to the preserver's custody. Both sets of requirements define and give a basis for assessing the records' identity and integrity, which must be preserved for the copies to be authentic.[3]

[1] See http://www.interpares.org/ip1/ip1_index.cfm.

[2] See Authenticity Task Force, "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records" in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 204–219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf. Abridged versions of the benchmark and baseline requirements are provided in Appendices 21a and 21b, respectively. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_21.pdf.

[3] As the InterPARES 1 Preservation Task Force determined, "[e]mpirically, it is not possible to preserve an electronic record: it is only possible to preserve the ability to reproduce the record. That is because it is not possible to store an electronic record in the documentary form in which it is capable of serving as a record. There is inevitably a substantial difference between the digital representation of the record in storage and the form in which it is presented for use." (Kenneth Thibodeau et al., "Part Three – Trusting to Time: Preserving Authentic Records in the Long Term: Preservation Task Force Report," in Duranti, *Long-term Preservation*, ibid., 106. Online reprint available at http://www.interpares.org/book/interpares_book_f_part3.pdf). In other words, only the first instantiation of a digital record, before it is stored, is an original. Once the first instantiation is saved, and thus stored in the system in the form of one or more digital components, the original record ceases to exist. Consequently, all subsequent manifestations of stored records are, *ipso facto*, copies. InterPARES 2 research expanded on this concept by distinguishing between a *stored* digital record—effectively defined as a digital object, placed in a storage system on a digital medium, that is managed as a record, and which includes information about the properties of the object and may also include methods of performing operations on or with the object—and a *manifested* digital record—effectively defined as a digital record that is visualized or rendered from a stored digital record and/or stored digital component(s) in a form suitable for presentation either to a person (i.e., in human-readable form) or to a computer system (i.e., in machine language) (see Luciana Duranti and Kenneth Thibodeau (2006), "The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES," *Archival Science* 6(1): 13–68. Note: A reprint of this article is included in Appendix 2.). In fact, "the primary

By focusing on the theoretical requirements for authenticity, the work of InterPARES 1 clarified obstacles to preservation in a thorough and coherent way, providing a complete framework in which to understand and resolve problems that many organizations have lately experienced. These problems were not simply the familiar results of degraded media and changing software. "For example, it highlighted the extent to which electronic systems are still being designed to manage data rather than records."[4] Case studies found that few systems contained entities satisfying the diplomatic definition of a record. Even systems that did contain records did not retain enough information about identity and integrity; so, by definition, the records could not be preserved authentically. The studies also encountered types of information displays that did not seem to have the fixity that one expects of records; for example, computer-monitor displays that assemble information from various, continuously updated sources. Like records, such displays inform the decisions and actions of organizations, but they are not stored or fixed, which raises the question of whether they could be preserved in any sense.

These findings call to mind problems of information management in activities far removed from business and law. One fifth of the data generated by the 1976 Viking exploration of Mars[5] and the works of nearly half of composers[6] and one-quarter of digital photographers[7] have been lost or threatened by technological obsolescence or inadequate preservation strategies. Challenges have been mounted to the trustworthiness of the records of electronic voting machines.[8] Every user of the Internet is familiar with broken hyperlinks, unplayable media and the challenge of determining whether information is authoritative and true.[9]

The nature of these activities gives hope that archival science can assist in finding solutions to these problems. Although artistic objects and experiences are not records in diplomatic theory (being final products, not by-products, of an activity), our appreciation of them generally requires knowledge of the actions and contexts in which they were produced.[10] Hypothesis-testing activities of science depend on the reproducibility of experiments, which in turn requires understanding exactly how recorded data were gathered and interpreted. And as governments mandate that their services be offered online, citizens will want their transactions mediated by interactive applications to be completely and accurately recorded in a way that allows them to trust the record.

But if these problems can indeed benefit from archival expertise, there are significant impediments to understanding them and to finding solutions through collaborations among

---

purpose of keeping the stored record is to be able to reproduce the manifest record, while the manifest record is preserved to communicate information to persons or other systems" (ibid., 51).

4 Heather MacNeil et al., "Part One – Establishing and Maintaining Trust in Electronic Records: Authenticity Task Force Report," in Duranti, *Long-term Preservation*, op. cit., 24. Available at http://www.interpares.org/book/interpares_book_d_part1.pdf.

5 See Terry Cook (1995), "It's Ten O'Clock, Do You Know Where Your Data Are?" *Technology Review* 98: 48–53; and Ross Harvey (2000), "An Amnesiac Society? Keeping Digital Data for Use in the Future." Paper presented at the LIANZA 2000 Conference, New Zealand, 15-18 October 2000.

6 Michael Longton (2004), "InterPARES 2 Project - General Study 04 Final Report: Recordkeeping Practices of Composers," 1. Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs04_final_report.pdf.

7 Jessica Bushey and Marta Braun (2006), "InterPARES 2 Project - General Study 07 Final Report: Survey of Recordkeeping Practices of Photographers using Digital Technology," 22. Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs07_final_report.pdf.

8 See, for example, http://www.votetrustusa.org/.

9 Chip Martel et al. (2001), "A General Model for Authentic Data Publication," 1. Available at http://www.cs.ucdavis.edu/~devanbu/files/model-paper.pdf; and Michael T. Goodrich et al. (2001), "Authenticated Data Structures for Graph and Geometric Searching," Technical report, Center for Geometric Computing, Brown University, 1. Available at http://www.cs.brown.edu/cgc/stms/papers/authDataStr.pdf.

10 David Davies, *Art as Performance* (Oxford: Blackwell, 2004).

archivists, creators and computer scientists. In the artistic, scientific and governmental sectors, the concepts of authenticity and reliability have diverse meanings, whose relation to those in archival science is not always clear. Also, the structure and function of the digital entities created in art and science do not always resemble those in administrative and legal contexts, so it is not clear how well the requirements established by InterPARES 1—or even the archival concept of authenticity itself—apply. Indeed, one of the most interesting research questions for InterPARES 2 was whether it is possible to satisfy these requirements for records (or other digital entities) in activities further removed from traditional recordkeeping practices. One way to answer this question was suggested by the Authenticity Task Force report,[11] which emphasized how important it is to study documents within the context of the systems in which they are created.

Accordingly, the second phase of InterPARES began in 2002 to develop a theoretical understanding of the records generated in interactive, experiential and dynamic systems, of their process of creation and of the present and potential use of records in the artistic, scientific and governmental sectors. The Project was organized into three domains of research, each tasked to investigate different aspects of the problem. The Domain 2 Task Force investigated the concepts of authenticity, reliability and accuracy as they are understood theoretically and practically in the artistic, scientific and governmental sectors, and it considered how those understandings relate to the InterPARES 1 definitions. To this end, the following three mandates were established for Domain 2 during the Project's February 2003 plenary workshop in Vancouver:

1. to find out how the concepts of reliability, accuracy and authenticity are used by records creators in each of the Project's three focus sectors;
2. to find out which words are used in each focus to signify these concepts; and
3. to find out what, if any, significance the creators in each focus place on these concepts.

Research team

The following is a list of researchers and research assistants who contributed to the work of the Domain 2 Task Force throughout the Project:[12]

Co-chairs:

Philip Eppard	Jan 2002 - Dec 2006
Brent Lee	Jan 2002 - Dec 2005
John Roeder	Jan 2002 - Dec 2006
Bill Underwood	Jan 2002 - Dec 2006

Researchers:

Marta Braun	Ryerson University, Canada—Working Group 2.1
Ann Butler	New York University, USA—Working Group 2.1
Hannelore Dekeyser	Katholieke Universiteit Leuven, Belgium—Working Group 2.3
Philip Eppard	University of Albany, State University of New York, USA—Working Group 2.3

[11] MacNeil et al., "Authenticity Task Force Report," op. cit., 24.
[12] Researcher membership in Domain 2 changed somewhat over the five years of the Project. Among those who were interested in Domain 2 issues but were unable to participate for the full length of the Project are: Margaret Campbell, Nova Scotia Provincial Archives, Canada; Ben Howell-Davis, Davis International Associates, USA; Reagan Moore, San Diego Supercomputer Center, USA; and Xiaowei Qiu, State Archives Administration of China.

Ken Hawkins	National Archives and Records Administrations, USA—Working Group 2.2
Ian Lancashire	University of Toronto, Canada—Working Group 2.1
Brent Lee	University of Windsor, Canada—Working Group 2.1
Michael Murphy	Ryerson University, Canada—Working Group 2.1
Eun G. Park	McGill University, Canada—Working Group 2.2
Richard Pearce-Moses	Arizona State Library—Working Group 2.3
John Roeder	The University of British Columbia—Working Group 2.1
Andrew Rodger	Library and Archives Canada—Working Group 2.1
Bill Underwood	Georgia Tech Research Institute, USA—Working Group 2.3

Research Assistants:

Scott Amort	The University of British Columbia, Canada
Gary Barclay	The University of British Columbia, Canada
Lindsey Bergen	The University of British Columbia, Canada
Natalie Catto	The University of British Columbia, Canada
Heather Dean	The University of British Columbia, Canada
Shanna Fraser	The University of British Columbia, Canada
Jessica Glidewell	The University of British Columbia, Canada
Joshua Hauck-Whealton	University at Albany, State University of New York, USA
Ted Hoppenstedt	University at Albany, State University of New York, USA
Tracey P. Lauriault	Carleton University, Canada
Rachel McMullin	University at Albany, State University of New York, USA
Peter Runge	University at Albany, State University of New York, USA
Mary Beth Sullivan	University at Albany, State University of New York, USA
Carol Ward	University at Albany, State University of New York, USA
Reginald White	University at Albany, State University of New York, USA
Mark Wolfe	University at Albany, State University of New York, USA
Catherine Yasui	The University of British Columbia, Canada
Sherry Xie	The University of British Columbia, Canada
Jessica Zacher	University at Albany, State University of New York, USA

Research Questions and Methodology

The goals of Domain 2 were articulated in the original Project proposal as a series of research questions.[13] In brief, these questions ask: What do the concepts of reliability, accuracy and authenticity mean in the context of artistic, scientific and governmental activities? To what extent, and how, do records creators in these areas presume and verify their records to have these qualities? How do these presumptions, if they exist, relate to the conceptual requirements for authenticity that the UBC-MAS[14] and InterPARES 1 projects generated for database systems? What intellectual tools, such as guidelines, and what technologies would assist creators to create authentic, reliable and accurate records while still respecting legal obligations, cultural differences, freedom of expression and inquiry and right to privacy?

[13] See Appendix 12.
[14] See http://www.interpares.org/UBCProject/index.htm.

One new direction, implicit in these questions, was that InterPARES 2 did not look only for instances of ideal digital records as they were modeled by InterPARES 1. The arts and science focuses seemed too different to warrant any such presupposition and, in any event, InterPARES 1 itself had discovered that even the documents in administrative systems were in many cases far removed from the ideal. Rather, as suggested by the Authenticity Task Force, the research considered all entities existing in actual systems as well as creators' conceptions of the nature, by-products and products of their activities, and what they understood to be required for presuming authenticity, reliability and accuracy.

Domain 2 accomplished its work through several avenues of investigation. First, researchers combed the literature specific to each focus (i.e., arts, science and government) to find discussions of authenticity, reliability and accuracy, and of related but differently named concepts. They then constructed and published on the InterPARES Web site annotated bibliographies. These bibliographies served as research tools for the other activities of the Project, which were manifested in papers and presentations about the conceptual analysis.

To ground these mostly theoretical discussions, researchers also analyzed reports generated by InterPARES studies of current practices in these fields—both studies of specific cases and more general surveys and interviews of creators. The studies were chosen as exemplars of current practice that showed the potential to cast light on conceptions of the authenticity and reliability of digital records. Domain 2 helped design the research instruments for these studies, in part, to elicit creators' views on its research questions.

The bibliographic research, conceptual analysis and case study reviews provided the main inputs to the products of the InterPARES 2 Terminology Cross-domain, an interdisciplinary research unit directed by lexicographers and experts in knowledge organization. Through a rigorous procedure, the Terminology Cross-domain developed a glossary, dictionary and ontology (a formal description of the concepts existing in the community of creators and preservers studied by the Project).[15] These products rationalized and controlled the language used by InterPARES 2 researchers, who came from quite various disciplinary and national traditions. Domain 2 research provided support, nuance and context to the definitions listed in the glossary as the official working concepts of the Project. For example, the glossary definition of "authenticity" as "the trustworthiness of a record as a record; i.e., the quality of a record that is what it purports to be and that is free from tampering or corruption" is enhanced and modified by the thirteen alternative definitions (listed in the dictionary) that Domain 2 located just in the arts.

At the outset of InterPARES 2, it was hoped that Domain 2 would describe a theory of how to make and keep the records of dynamic and interactive systems in a way that would consider their diverse cultural and disciplinary environments. Progress was made towards such a theory: researchers proposed expansions to the traditional conceptions of a record and of metadata that were appropriate to the interactive and dynamic environments that the Project studied; and sets of guidelines for records creation and maintenance were developed to address the various problems discussed in the literature and observed in the case studies.

Lastly, Domain 2 initiated a test project to transform the documents in one of the case studies into preservably reliable, accurate and authentic records. This was informed not only by the theoretical investigations outlined above, but also by the Domain's participation in modeling sessions that identified the procedures, inputs, outputs, resources and controls on creation.

This report summarizes the results of Domain 2's work. Each of the following three sections addresses one of the InterPARES 2 Focus Task Forces. Within each section, the conceptual

[15] See http://www.interpares.org/ip2/ip2_terminology_db.cfm.

analysis is reviewed and compared with the analyses of the case studies and general studies relevant to that focus. The concluding three sections of this report consider the extent to which the tools developed by InterPARES 1 are adequate to cover the preservation concerns of the focuses, and set forth the basis for guidelines to assist creators in producing preservably authentic records, should they so desire.

At the most general level, Domain 2's findings are not surprising. For the most part, creators' conceptions conform to the various senses of the terms exposed in the theoretical literature of their discipline (for example, by those authors surveyed in the review of these concepts in the arts, summarized below); however, they do so in informal and overlapping ways. Artists, scientists and bureaucrats have very different ideas about the documents they create and reference, what needs to be kept and the features that are essential; terms that have a fairly precise meaning to the archival profession have very different, even contradictory, meanings to these creators. The diversity the Domain 2 researchers found has been reflected in some current thinking about the constructed nature of the concept of authenticity.[16] It is hoped that the specifics of the conceptual analysis presented here, which attempts to carve out semantic boundaries and make clear distinctions among similarly-named concepts, will promote better communication among all interested parties.

The Domain 2 researchers found that although the case studies were quite diverse, they shared many common problems: technological obsolescence, lack of control over creation procedures, insufficient documentation and uncertainty about what digital objects needed to be saved. The ubiquity of the problems identified helped focus the drafting of guidelines for making and maintaining digital materials.[17]

The Domain 2 researchers also found that the benchmark requirements of InterPARES 1 were indeed useful for measuring a presumption of the authenticity of creators' records, for many instances were observed in which documents could not be preserved because they lacked some essential attribute that the requirements identified. However, the researchers also found that it is difficult to apply, or even to adapt, the requirements to the variety of systems that the Project studied, and many cases were observed where they were not sufficient to preserve the kinds of authenticity that the creators valued. The conceptual analyses and experiences of the case studies provide valuable insight into these additional aspects of preservation.

Focus 3 – Government

Scope of the research

The governmental sector presented a terrain that was in many ways very familiar to InterPARES researchers. Indeed, the bulk of the work done studying issues around the creation, maintenance and preservation of digital records has been on the records of governments or similar bureaucratic entities. Yet, InterPARES 2 concentrated its focus on records in interactive, dynamic and experiential systems, particularly those records created in what has come to be called e-government—the use of digital technology to improve the delivery of governmental information and services to the citizenry. Such services typically delivered in an interactive mode through the World Wide Web pose new challenges to both creators and preservers. But the

[16] Heather MacNeil and Bonnie Mak (2007), "Constructions of Authenticity," *Library Trends: Recent Trends in CulturalHeritage Preservation* 56(1): 26–52.
[17] See Appendix 20.

constraints on the creators of government records result in a different environment than the ones studied by the other two focus areas of InterPARES 2. The freedom of expression enjoyed by records creators in the arts, for example, is not at all characteristic of the bureaucratic laws and regulations controlling the e-government environment. Accountability is the watchword with government records as public officials strive to ensure the rights of citizens while maintaining the ability to demonstrate their own faithful execution of duties in the public trust.

Conceptual analysis: authenticity, accuracy and reliability in the literature of e-government

The literature review studying the concepts of authenticity, accuracy, and reliability in e-government rather quickly shows that these concepts are seldom addressed directly. Most e-government literature is focused on delivery options and how to improve them—in short, obtaining the maximum efficiency in the use of information technology to meet the needs of citizens. Any concerns about authenticity in the electronic environment are generic, without singling out the records produced by these newer kinds of electronic systems. This is understandable, for consistency across records in different formats is appropriate. At the same time, it seems clear that older concerns about such issues have carried over from the paper environment. Of these three concepts addressed in Domain 2, it is authenticity that has received the most attention within the governmental sector.

A good example of how these issues have been discussed in the governmental sector can be seen in the electronic records guidelines issued by the New York State Office for Technology. The guidelines were developed "to provide general direction on how state and local government agencies can ensure the authenticity, integrity, security, and accessibility of electronic records (e-records)."[18]

Archivists are keenly aware of how terminology in their field has been used imprecisely, and this is a particular problem in the area of digital records.[19] It should not be surprising, therefore, that the terminology used in the governmental sector in discussing digital records is at times vague or inconsistent in its usage. This is particularly true for words like "authenticity," "accuracy" and "reliability," which are not technical terms at all, but words with common sense, everyday meanings.

The New York State guidelines include a glossary, thereby clearly acknowledging the need to define some of the technical terms used in the document as well as those more common words in need of a precise definition. "Authenticity" is included in the glossary but accuracy and reliability are not. The definition of authenticity, however, is very constricted: "[It] refers to the methods used to verify the source or origin of an e-record. Authenticity is closely related to the concept of *integrity*."[20] The InterPARES definition of authenticity is more in line with the concept of integrity defined in the New York State glossary, which begins by asserting that integrity "is the attribute that the record's contents have not been changed, deleted or otherwise altered."[21] This definition goes on to draw in accuracy as part of integrity, stating that "[i]n addition, integrity addresses the accuracy and timeliness of the contents of a record." Finally, this

[18] New York State Office for Technology (2002), "E-Records Guidelines: Ensuring the Security, Authenticity, Integrity, and Accessibility of Electronic Records." Part 4 of *Electronic Signature and Records Act (ESRA) Guidelines*. Available at http://www.oft.state.ny.us/arcPolicy/policy/ESRAGuidelines4.htm.
[19] This is one of the reasons why InterPARES 2 established its Terminology Cross-domain Task Force.
[20] New York State Office for Technology (2002), "Glossary." Part 5 of *Electronic Signature and Records Act (ESRA) Guidelines*. Available at http://www.oft.state.ny.us/arcPolicy/policy/ESRAGuidelines5.htm. Italics as in original.
[21] Ibid.

definition emphasizes the legal importance of maintaining authenticity and integrity, noting that "[b]oth authenticity and integrity are derived from the legal arena and have a strong bearing on the legal admissibility of records."

The fact that the accuracy of the content of a record should be subsumed under integrity is a bit surprising, since accuracy of content is not addressed in the ESRA guidelines themselves. Instead, accuracy is used as an attribute of systems: "The reliability and accuracy of the systems, processes and procedures used to create, capture, and maintain e-records are critical to demonstrating their authenticity and integrity."[22] Similarly, later in the guidelines, government workers are cautioned: "Make sure the system performs in an accurate, reliable, and consistent manner in the normal course of business" to ensure that records are acceptable "for legal, audit, and other purposes."[23] Thus, accuracy is addressed only peripherally, and generally as an attribute of the way a system should function. Otherwise, it might be assumed that maintaining the integrity of a record will also ensure that the accuracy of the contents of the record are also maintained intact.

Similarly, "reliability" is a concept that seems more to be a characteristic of systems than of records, as noted above in the quote about reliability and accuracy of systems being central to demonstrating the authenticity and integrity of digital records. This is expressed elsewhere in the ESRA guidelines more specifically as "the reliability of hardware and software" affecting "the authenticity and integrity of e-records."[24] In sum, then, accuracy and reliability are viewed primarily as attributes of systems or the functioning of systems. It seems to be implicit that reliable systems will keep reliable records. Perhaps the accuracy of the content of records is not addressed more specifically because its importance is taken as a given that is not unique to the world of digital records. Such a position, however, overlooks the greater potential for change in the content of records in a digital environment.

A more fruitful discussion can be had of the term "authenticity" and its closely allied term "integrity." The ESRA glossary definition clearly establishes a link between authenticity and authenticate when it says that authenticity "refers to the methods used to verify the source or origin of an e-record." The guidelines themselves, however, almost always link authenticity to integrity; indeed, they seem to be virtually synonymous.

Preference for use of the word "integrity" instead of "authenticity" is evident in the United States *E-Government Act of 2002*.[25] Here integrity encompasses authenticity, for it is defined as "guarding against improper information modification or destruction, and includes ensuring information nonrepudiation and authenticity,"[26] where nonrepudiation is defined as "the ability to ensure that a party to a contract or a communication cannot deny the authenticity of their signature on a document or the sending of a message that they originated."[27] Integrity is linked with confidentiality and access as part of information security, which is applicable to "protecting information and information systems."[28] "Accuracy" is referenced in the Act in the context of how the integration of federal information systems will help assure and validate the accuracy of information. Obviously, accuracy is regarded as important, but it is assumed as a given requirement for the content of records and thus is not given much direct attention here or

[22] Ibid., Part 4, op. cit.
[23] Ibid.
[24] Ibid.
[25] United States Congress (2002), *E-Government Act of 2002*. 107th Cong., 2d sess., 2002. H.R.2458.ENR.
[26] Ibid., Sec. 3542(b)(1)(A).
[27] SearchSecurity.com. Available at http://searchsecurity.techtarget.com/sDefinition/0,,sid14_gci761640,00.html.
[28] United States Congress, *E-Government Act of 2002*, op. cit., Sec. 3542(b)(1).

elsewhere in the literature. "Reliability" is a term not used in the Act, although one might infer that it would be considered as part of the discussion of information security.

Terminology and definitions found in documents on digital records and data on various U.S. state Web sites show a similar range of imprecision in application of these terms, though all are within a commonly accepted sense. The word "trustworthy" has been adopted by the state of Minnesota in its *Trustworthy Information Systems Handbook*, and trustworthy is defined as an attribute of records that "contain information that is reliable and authentic."[29] Clearly, the title indicates that trustworthy applies to systems as well as to records themselves. The handbook's glossary formally says, "Authenticity is a function of a record's preservation and is a measure of a record's reliability over time," while reliability is defined as "the measure of a record's authority and is determined solely by the circumstances of the record's creation." The state of Texas offers a definition of authenticity that is similar to that proposed by InterPARES, but reliability seems to refer to the ability to sustain and reproduce records accurately into the future.[30] The state of Wisconsin offers a similar definition of "reliable," while its definition of "authentic" is that "the retained electronic record correctly reflects the creator's input and can be substantiated."[31] Unlike most such documents, the Wisconsin standards also define "accurate" as meaning that "all information produced exhibits a high degree of legibility and readability and correctly reflects the original record when displayed on a retrieval device or reproduced on paper."

Authenticity, accuracy and reliability in the governmental sector case studies

InterPARES 2 researchers carried out eight case studies within the government focus area. The case study reports provide additional insight into how those working in the governmental sector regard the concepts of authenticity, reliability and accuracy in the context of their own electronic systems. The reports tend to confirm much of what the conceptual analysis shows about authenticity, reliability and accuracy in the e-government area. Often, authenticity is either presumed, providing that requisite procedures have been followed, or is tied to authentication methods, such as PKI (Public Key Infrastructure). There is concern with accuracy of information, but it is likely to be limited to the accuracy of data at the point of creation, after which accuracy (like authenticity) is presumed to be protected by procedural controls.

One of the more useful discussions of these issues in the case study reports is in case study 20, Ireland's Revenue On-Line Service (ROS). Here there is a presumption of the "authenticity of received, signed and submitted tax forms from authorised users."[32] In addition, a "chain of authenticity" is presumed based on "user log-ins, digital certificates and PKI."[33] This "security wrapper" is retained "to confer authenticity and non-repudiation over time."[34] The study acknowledges that it is unclear whether this approach will work over time, however. Nevertheless, it is clear that the ROS relies on these external controls to convey authenticity and

[29] Minnesota Historical Society, State Archives Department (2002), *Trustworthy Information Systems Handbook*. Version 4. Available at http://www.mnhs.org/preserve/records/tis/tableofcontents.html.
[30] Texas Department of Information Resources (2004), *Architecture Components for the Enterprise*. *Data and E-Records Management Domain*. Available at http://www.dir.state.tx.us/ace/documents/phase1toc.htm.
[31] Wisconsin Department of Administration (2001), *Administrative Rules: Adm 12, Electronic Records Management - Standards and Requirements*. Available at http://www.legis.state.wi.us/rsb/code/adm/adm.html.
[32] John McDonough, Ken Hannigan and Tom Quinlan (2005), "InterPARES 2 Project - Case Study 20 Final Report: Revenue On-Line Service (ROS)," 77. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs20_final_report.pdf.
[33] Ibid.
[34] Ibid.

sees this authenticity as carrying through time. The discussion of reliability in this case study suggests that it is a concept not well differentiated from authenticity. Again, it is the controlled environment, through the use of passwords and PKI, which confers reliability on records. In this case study, accuracy relates to the factual accuracy of the data in records, but since individuals and revenue agents may enter inaccurate data, full accuracy of the data cannot be guaranteed. "A certain number of business rules and logic to check calculations" are built into the system to help guard against inaccurate data being entered.[35]

In the City of Vancouver Geographic Information System (VanMap) case study (case study 24), accuracy is sometimes referred to as "data quality." There is a concern that information added to VanMap be as accurate as possible before it is entered. Since data comes from external sources, however, its accuracy cannot be guaranteed. However, the VanMap team does guarantee that information is as accurate as it was when it was entered into the system, that "the data are not altered in such a manner as to affect their accuracy and reliability."[36] Note that in this answer reliability is treated as a characteristic of data. When decisions based on information in VanMap are being made, staff may verify the data or seek other kinds of independent verification. Clearly there is a concern for the factual accuracy of data in the system, since there is a recognition that it can affect decision-making. In fact, this case study treats authenticity as residing largely in the accuracy of the data.

Case study 21 looks at the electronic filing system in place at the Supreme Court of Singapore. In this case study, "reliability" at first seems to refer to the overall system, for it relates to the ability of the court to process submissions from attorneys and litigants in an efficient manner and to keep track of the large number of files in the system. But the records are considered reliable as well, because "they are created and modified in a controlled environment with access privileges assigned to respective action officers based on their job responsibilities."[37]

Accuracy in the Singapore case relates to "the provision of accurate information from the case records."[38] Various methods are in place to ensure that the records are accurate, although the concern with accuracy seems to be centred more on creation of the records than on their maintenance. Perhaps this is seen more as a matter of protecting authenticity, for, once the records are created, preserving their authenticity would seem to include preserving accuracy. Here one sees also the confusion of authenticity and authentication. The use of authentication technologies is the source of a presumption of authenticity. By use of PKI and other security controls, alteration or tampering of information in the files is prevented and authenticity is ensured, although not necessarily for the long term.

Of all the case studies in the government focus, the computerization of the Alsace-Moselle land registry (case study 18) offers perhaps the most sophisticated technological awareness of issues of accuracy, authenticity and reliability. For more than a century, a rigorous system was in place to provide "accurate, reliable, and authentic information" through a paper recordkeeping system that has now been computerized.[39] The case study report details all the steps taken in development of the new system. Here it is only important to pay attention to how the concepts of

[35] Ibid., 75.
[36] Evelyn McLellan (2005), "InterPARES 2 Project - Case Study 24 Final Report: City of Vancouver Geographic Information System (VanMap)," 26. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs24_final_report.pdf.
[37] Elaine Goh (2005), "InterPARES 2 Project - Case Study 21 Final Report: The Electronic Filing System (EFS) of the Supreme Court of Singapore," 39. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs21_final_report.pdf.
[38] Ibid.
[39] Jean-François Blanchette, François Banat-Berger and Geneviève Shepherd (2004), "InterPARES 2 Project - Case Study 18 Final Report: Computerization of Alsace-Moselle's Land Registry," 20. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs18_final_report.pdf.

accuracy, authenticity and reliability are viewed in the process. First, one can note that quality of the data is the way in which the concept of accuracy is expressed. Authenticity, integrity and reliability are all regarded as attributes of data or information, and the terms are often used interchangeably. Authentication techniques are key features of the system and are seen as central to the guarantee of authenticity, integrity or reliability.

A somewhat different take on the concept of accuracy emerges in the case study of the Archives of Ontario Web Exhibits (case study 05).[40] Here accuracy is seen not as relating to the digital components of the records so much as to the concept of historical or narrative accuracy of an exhibit as a whole. In other words, the creators of the Web exhibits seek to present factual information with proper documentation but leave interpretation to those viewing the exhibit. Although this seems to be a significantly different conception of accuracy in digital records, one could argue that this is analogous to the concern for correct information being entered into tax forms in the ROS case study noted above. In case study 05, authenticity is presumed, particularly within the environment of the creator, but whether these Web exhibits can be considered to be "authentic" when viewed by users will depend in part on how they are displayed in any given technological environment.

This sampling of findings on authenticity, reliability and accuracy from the InterPARES 2 case studies in the government focus indicates that there is an awareness of the importance of these three concepts. Reliability is less often addressed, and when it is it seems to be used as a synonym for authenticity or at least is seen as inextricably connected with authenticity. Both authenticity and reliability are deemed to be protected by procedural controls and authentication techniques. Accuracy is sometimes equated with data quality, a somewhat fuzzier concept, and one which was not detected in the literature review in the government focus. The controlled bureaucratic process in government lends a sense of confidence that structure and procedure will help maintain these essential characteristics of government records.

Conclusions

Because of the long-recognized necessity for government to be accountable for its actions, the need to maintain records that can be demonstrated to be authentic and reliable is not a new concept for those responsible for creating and maintaining government records. As Minnesota's *Trustworthy Information Systems Handbook* states, "We need trustworthy information systems to ensure our accountability as government agencies."[41] Reliability and authenticity are characteristics of information and records that are essential to trustworthiness. The fact that so much attention has been paid to the challenges posed by digital records in the governmental sector is evidence that concern for these issues has taken strong root. Government archives and records managers have taken the lead in impressing these issues on executives, agency heads and information technology personnel. The InterPARES research in Domain 2, however, suggests that terminology is used loosely and that adjectives such as "authentic," "trustworthy" and "reliable" are applied at different times to information, data, records and systems. Accuracy is either considered a characteristic of authenticity or something that is generally beyond the control of government workers, insofar as it relates to the factual accuracy of data entered into forms by others. As with paper records, authenticity is often presumed, particularly in instances

[40] Jim Suderman et al. (2004), "InterPARES 2 Project - Case Study 05 Final Report: Archives of Ontario Web Exhibits." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs05_final_report.pdf.
[41] Minnesota Historical Society, *Trustworthy Information Systems Handbook*, op. cit.

where authentication techniques are employed. Although there is a commendable respect for authenticity, reliability and accuracy in the government focus, there needs to be a greater sense of the complexities involved in maintaining these characteristics for records requiring long-term preservation.

Focus 2 – the Sciences[42]

Scope of the research

The crucial importance to society of scientific research and its ready use of the latest technologies were key factors in leading InterPARES 2 researchers to make scientific activities one of the Project's three focuses for investigation. Archivists have done significant work in the area of documentation and appraisal in the sciences—work that has been supplemented by the scientific community's efforts in recent years[43]—to address the impact of information technologies on scientific research and recordkeeping. With traditional archival definitions of the concept of record under review in the digital era, scientific activities also seemed to be a fruitful area of study because of the different perspectives that scientific researchers have about records and recordkeeping practices.[44]

The sciences are a broad and heterogeneous area for study, and it was not possible for InterPARES researchers to investigate all branches of the world of science. The study of the scientific concepts of authenticity, accuracy and reliability, however, began with a broad literature review across the disciplines, seeking specific discussions of, or references to, those three key concepts. The findings of the literature review were supplemented with data from various InterPARES 2 case studies that focused on scientific data and records creation, management, appraisal and retention. Finally, broader, more disciple-wide perspectives on these issues were provided by a number of general studies carried out by the Focus 2 researchers.

---

[42] InterPARES wishes to acknowledge SUNY Graduate Research Assistant, Joshua Hauck-Whealton, for his help in compiling the case study and bibliographic data used in this section.

[43] See, for example, National Research Council, Commission on Physical Sciences, Mathematics and Applications, *Preserving Scientific Data on Our Physical Universe: A New Strategy for Archiving the Nation's Scientific Information Resources* (Washington, D.C.: National Academy Press, 1995). Available at http://www.nap.edu/catalog.php?record_id=4871#toc; National Science Foundation, *Report of the National Science Board: Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century*, NSB-05-40, September 2005. Available at http://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf; David F. Strong and Peter B. Leach, *National Consultation on Access to Scientific Data: Final Report* (Ottawa: Canadian Institute for Scientific and Technical Information, National Research Council Canada, 2005). Available at http://ncasrd-cnadrs.scitech.gc.ca/NCASRDReport_e.pdf; Kenneth Thibodeau (1995), "Preserving Scientific Data on Our Physical Universe," *IASSIST Quaterly* 19(4): 26–29. Available at http://iassistdata.org/publications/iq/iq19/iqvol194thibodeau.pdf; Joan Warnow-Blewett, Joel Genuth and Spencer R. Weart, *AIP Study of Multi-Institutional Collaborations: Final Report. Highlights and Project Documentations* (College Park, MD: Center for History of Physics, American Institute of Physics, 2001). Available at http://www.aip.org/history/pubs/collabs/highlights.pdf; Library of Congress, *National Digital Information Infrastructure and Preservation Program (NDIIPP), Digital Preservation*. Available at http://www.digitalpreservation.gov/index.html; David L. Brown, Grace Welch and Christine Cullingworth (2005), "Archiving, Management and Preservation of Geospatial Data: Summary Report and Recommendations," GeoConnections Policy Advisory Node: Working Group on Archiving and Preserving Geospatial Data. Available at http://www.geoconnections.org/publications/policyDocs/keyDocs/geospatial_data_mgt_summary_report_20050208_E.pdf; and CODATA Working Group on Archiving Scientific Data, The Committee on Data for Science and Technology (CODATA) of the International Council for Science. Available at http://www.nrf.ac.za/codata/.

[44] See Tracey P. Lauriault, Barbara L. Craig, D. R. Fraser Taylor and Peter L. Pulsifer (2007), "Today's Data are Part of Tomorrow's Research: Archival Issues in the Sciences," *Archivaria* 64 (Fall): 123–179.

Conceptual analysis: authenticity, accuracy and reliability in the literature of the sciences

A recent European Task Force on permanent access to scientific records neatly lays out the scope of the definitional problem:

> The definition of "the records of science" must be broad to ensure an accurate record of the research process and its results is created. Within this definition it is essential to include both the formal, structured 'minutes of science' (published records in the formal refereed scientific literature) and the less structured, informal communication mechanisms which are now commonly used to share topical ideas and information (such as Web sites, moderated bulletin boards and email). In addition, through improvements in computing and networks linked with powerful data and text mining techniques, new research practices have developed which are data-intensive and highly collaborative. Scientific records may take the form of raw data, aggregated into datasets relevant to a particular topic or field. These can be numeric, graphic or textual and may contain embedded logic (chemical compound structures, crystallographic data, genome sequences).[45]

This wide-ranging definition includes the published results of scientific research, which would not be considered "records" in the traditional archival sense, as well as forms of documentary communication, such as e-mail, readily recognizable as records. But the real emphasis in this definition seems to be on scientific data and the various aggregations in which they are gathered and utilized for research purposes. InterPARES 2 research into scientific activities confirms that most scientists do not think of records in the archival sense but rather tend to equate "scientific records" with "scientific data," sometimes using the term "scientific data records." Thus, it is important for archivists to understand that for scientists, scientific data can also be defined more precisely to mean "numerical quantities or other factual attributes generated by scientists and derived during the research process (through observations, experiments, calculations and analysis),"[46] and "numbers, images, video or audio streams, software and software versioning information, algorithms, equations, animations, or models/simulations."[47] Although the former definition is more in keeping with the definition of *data* provided in some archival contexts,[48] it is interesting to note that the latter definition circumscribes a broader concept that includes digital entities that, depending on their context of creation and use, archivists would more readily identify as records.

Nevertheless, the Focus 2 research suggests that, in the eyes of many scientists, maintaining archives in the sciences is less an act of recordkeeping and more an act of data management and processing. "A key component of creating the public archive of information is the efficient capture and curation of the data—data processing."[49] Parallel to this difference in focus is the

---

[45] European Task Force Permanent Access (2005), "Permanent Access to the Records of Science: Proposal for a Research & Development Programme," 3–4. Available at http://www.alliancepermanentaccess.eu/Proposal%20Research%20and%20Development.doc.
[46] CODATA Working Group on Archiving Scientific Data, op. cit.
[47] National Science Foundation, *Report of the National Science Board*, op. cit., 18. Available at http://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf.
[48] For example, the word *data* is defined as "Facts, ideas, or discrete pieces of information, especially when in the form originally collected and unanalyzed," by Richard Pearce-Moses in his *A Glossary of Archival and Records Terminology* (Chicago: Society of American Archivists, 2005). Available at http://www.archivists.org/glossary/.
[49] Helen M. Berman et al. (2000), "The Protein Data Bank," *Nucleic Acids Research* 28(1): 235–242. Available at http://nar.oxfordjournals.org/cgi/content/full/28/1/235.

perception that the scientific community evinces less overt interest in authenticity than it does in accuracy and reliability.

Of the three concepts of authenticity, reliability and accuracy, it is the last concept, accuracy, that receives the greatest attention in scientific literature; more specifically, concepts related to data quality. This is no doubt attributable to the fact that scientists focus on data, not records, and the accuracy of data is obviously crucial to the validity of scientific research. Conversely, the concept of accuracy is not as prominent in archival theory, but the InterPARES definition of accuracy is broad enough to include data and datasets as well as documents and records. Data are different than records, of course. Data are often still in the process of being used and modified, and have thus not been properly "set aside" in the fashion of records.

In an archival context, an accurate record is one that contains correct, precise and exact data, which is often adjudged in relation to the absoluteness of the data it reports or its perfect or exclusive pertinence to the matter in question. Furthermore, in an archival context, the accuracy of a record is *assumed* when the record is created and used in the course of business processes to carry out business functions, based on the assumption that inaccurate records harm business interests.[50] This assumptive approach toward the assessment of accuracy stands in marked contrast to the approach used in scientific research, where, because errors and uncertainty are a given, a general analytical tenant is that "no number has meaning unless it is accompanied by an estimate of uncertainty."[51] It is for this very reason, in fact, that accuracy is seen by most scientists as the most common and critical metadata element.

To an archivist, an authentic record does not have to be an accurate record, however. Although it is true that an authentic record is as reliable and accurate as it was when first generated, this is not the same thing as saying that authenticity ensures that the content of a record at the point of its creation is accurate. Thus, authenticity alone does not "automatically imply that the content of a record is reliable"[52] or accurate. Scientists, on the other hand, give primacy to data quality, which includes the concept of authenticity, normally articulated as data provenance or lineage. In this context, data accuracy is critical and the data need to be reliable. Data quality is normally articulated in a dataset's metadata; without metadata or data quality parameters, a scientist will not use, trust or rely on those data. Although each scientific discipline differs in how it defines scientific data quality, most include some or most of the following data quality elements: positional accuracy; attribute and thematic accuracy; completeness; semantic accuracy; and temporal information, reliability, lineage, logical consistency and objectivity.[53] In the fields of Geography and Geomatics, a number of standards have been published regarding the quality of geographic data.[54] The procedures described in the International Standard, ISO 19113:2002, provide a consistent and standard manner to determine and report a dataset's quality. The International Cartographic Association has also written a book on the subject, entitled *Elements of Data Quality*.[55]

The concept of precision is related but distinct from accuracy. "*Precision* refers to how *exact* and *reproducible* a measurement or estimate is, irrespective of its accuracy, while *accuracy*

---

[50] See *Creator Guidelines*, guideline 4, in Appendix 20.
[51] National Research Council, *Preserving Scientific Data*, op. cit., 37.
[52] Pearce-Moses, *A Glossary of Archival and Records Terminology*, op. cit.
[53] Stephen C. Guptill and Joel L. Morrison (eds.), *Elements of Data Quality* (Oxford: Elsevier Science, 1995).
[54] International Organization for Standardization, ISO 19113:2002 - Geographic Information—Quality Principles; International Organization for Standardization , ISO 19114:2003 - Geographic Information—Quality evaluation procedures.
[55] Guptill and Morrison, *Elements of Data Quality*, op. cit.

refers to how *close* a measurement or estimate is to the correct value."[56] A device used for measurement may be precise, by returning measurements that are always the same or close to the same, but not be accurate. A dataset may be reasonably accurate when all data points are close to the actual reality, but fail to be precise because of a wide spread of data. Although the InterPARES definition of accuracy encompasses the concepts of correctness and precision, the two concepts have different and important meanings within the sciences.

An additional feature of the concept of accuracy in the sciences is how it is affected by the idea of timeliness. It is generally assumed that there will be advances in the methods of data collection, and thus "evolving, improving accuracy of the determination" of data.[57] As a result, a more recent measurement will be considered more accurate; that is, more correct or closer to the actual reality. This suggests that data have a kind of shelf life, since the creation of a new (and more accurate) dataset will render an older dataset obsolete. It is recommended that obsolete data sets be discarded: "If the data have been completely superseded by better data . . . destruction of the records may be in order."[58]

It is useful here to understand the difference between this idea of data shelf life and another concept referred to in the literature as "currency" or "temporal accuracy." Both concepts suggest that older data should be replaced by more current data. However, the idea of currency reflects the fact that the reality itself has changed and that the presence of more accurate data is not just the result of improved methods or instruments. Currency, therefore, applies in situations where the facts being measured shift and change, such as in the compilation of maps or in population surveys. By the strictest definitions, data that no longer have currency are still accurate, as they truthfully depict reality the way it was at the time the measurement was taken. However, they may no longer be useful to the scientist for whom temporal accuracy is crucial.

As indicated earlier, the concept of authenticity does not loom large in the scientific literature. Any concern for authenticity that occurs is often part of the greater concern for accuracy. Several sources stress the need to properly assign responsibility for a given source of data so that errors can be fixed through correspondence with the creators. An article on digital signatures in electronic health records asserts, "The attribution of responsibility is a means to ensure the accurateness of the information."[59] In a paper on the protein data bank, the author writes, "In almost all cases, serious errors detected by these checks are corrected through annotation and correspondence with the authors."[60]

Ironically, however, although most organizations aim to ensure that their data are accurate, reliable and authentic, the Focus 2 case and general studies observed that many of these same organizations add disclaimers to absolve themselves of any responsibility for damages that may result from the use of their data. These types of disclaimers were used by a number of the scientific data portals surveyed in general study 10. Examples include the Antarctic Digital Database (ADD) (data portal IP2SF27), which cautions that its maps, when combined, may

---

[56] Randy Preston (2006), "InterPARES 2 Project - General Study 09 Final Report: Digital Recordkeeping Practices of GIS Archaeologists Worldwide: Results of a Web-based Survey," 74. Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs09_final_report.pdf. Italics as in original.

[57] National Research Council, Committee on Issues in the Transborder Flow of Scientific Data. *Bits of Power: Issues in Global Access to Scientific Data* (Washington, D.C.: National Academy Press, 1997), 48. Available at http://www.nap.edu/catalog.php?record_id=5504#toc.

[58] John L. Faundeen (2003), "The Challenge of Archiving and Preserving Remotely Sensed Data," *Data Science Journal* 2: 162. Available at http://www.jstage.jst.go.jp/article/dsj/2/0/159/_pdf.

[59] J. J. Bos (1996), "Digital Signatures and the Electronic Health Records: Providing Legal and Security Guarantees," *International Journal of Bio-Medical Computing* 42(1-2): 159.

[60] Berman et al., "The Protein Data Bank," op. cit., 237.

reflect some inconsistencies, particularly when older datasets are included; the National Geophysical Data Center (data portal IP2SF26) and the World Data Center for Solar Terrestrial Physics (data portal IP2SF10), both of which indicate that the Government of the United States and its employees cannot be held accountable for any data quality warranties; the FMRI Data Center (data portal IP2SF7), which absolves itself from liability in relation to data quality; the Indiana University Bio Archive (data portal IP2SF5), which reminds users that data contain errors; and the British Atmospheric Data Centre (BADC) (data portal IP2SF1), which absolves itself from responsibility of data on its Web site and once downloaded onto the user's computer.[61] A slightly different type of disclaimer is provided by the Cybercartographic Atlas (case study 06), which states that its content is intended to be used primarily for educational purposes and that the "[d]isclaimers and caveats on the Web site are intended to limit the [legal] responsibility of the Creator."[62]

These examples suggest that although scientists are unlikely to use the word authenticity, they are nevertheless very concerned about the identity and integrity of the data they use. Identity and integrity are the key components of authenticity in the understanding of InterPARES.[63] Within a science context, identity is established by metadata that record the phenomena observed and the kinds of measurements obtained and by which instruments and at what time and place and under whose responsibility. Integrity, on the other hand, means that the data have not been altered, either by unauthorized tampering or by corruption due to technical failure since they were first created. In the sciences, data integrity is often guarded first by *authentication* and other security measures to prevent unauthorized tampering, then by checksums or other techniques to spot altered bits.

As used in the field of Communications Science and Engineering, *authentication* refers to "security measures designed to protect a communication system against fraudulent transmissions and establish the authenticity of a message," while an *authenticator* is a "letter, numeral or groups of letters attesting to the authenticity of a message or transmission."[64]

As in other areas, however, the process of authentication and the use of authentication technologies are not sufficient in themselves to be a guarantor of authenticity of preserved scientific records over time, since authentication can, at best, only establish the authenticity of a record at *a specific point in time*. Moreover, authentication may, in many instances, simply mean establishing the identity (and therefore the trustworthiness) of an agent associated with a record, not of the data or record itself: "Authentication is a security service that consists of verifying that someone's identity is as claimed."[65] On the other hand, authentication procedures and technologies can be an important and perhaps, in some cases, essential element in an overall preservation strategy that is designed to guarantee the authenticity of preserved scientific records over time.[66]

---

[61] See Tracey P. Lauriault and Barbara L. Craig (2007), "InterPARES 2 Project - General Study 10 Final Report: Preservation Practices of Scientific Data Portals." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs10_final_report.pdf.
[62] Tracey P. Lauriault and Yvette Hackett (2005), "InterPARES 2 Project - Case Study 06 Final Report: Cybercartographic Atlas of Antarctica," 5, 27. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs06_final_report.zip.
[63] See MacNeil et al., "Authenticity Task Force Report," op. cit., 47, specifically, the section titled, "Conceptual findings: the requirements for authenticity."
[64] McGraw-Hill, *Dictionary of Scientific and Technical Terms*, sixth edition (New York: McGraw-Hill, 2003).
[65] Audun Jøsang and Mary Anne Patton, "User Interface Requirements for Authentication of Communication," in *Proceedings of the Fourth Australian User interface Conference on User interfaces 2003 - Volume 18*, R. Biddle and B. Thomas, eds. ACM International Conference Proceeding Series, vol. 36 (Darlinghurst, Australia: Australian Computer Society, 2003), 75. Available at http://portal.acm.org/ft_gateway.cfm?id=820105&type=pdf&coll=&dl=&CFID=15151515&CFTOKEN=6184618.
[66] See, for example, William E. Underwood (2002), "A Formal Method for Analyzing the Authenticity Properties of Procedures for Preserving Digital Records," in *Proceedings of the 2002 International Conference on Digital Archive Technologies*

Within the sciences, the trustworthiness of data is also conceived in terms of "data quality," which, although it includes the concept of authenticity, is normally articulated as data provenance or lineage. Data quality is normally articulated in a dataset's metadata; without metadata or data quality parameters, a scientist will not use, trust or rely on those data. Each scientific discipline differs in how it defines scientific data quality, as is demonstrated from the results of the Focus 2 case and general studies. However, most disciplines include some or most of the following data quality elements: positional accuracy; attribute and thematic accuracy; completeness; semantic accuracy; and temporal information, reliability, lineage, logical consistency and objectivity.[67]

Data "lineage" is information about the chain of transmission, from the moment the data were originally recorded, that brought the data to the user. Lineage speaks to the history of a dataset, its lifecycle from data collection to its many stages of compilations, corrections, conversions and transformations and the generation of new interpreted products. This concept might also be characterized as "data provenance" and is clearly related to data integrity. As with traditional paper records, the provenance of a particular scientific dataset is essential in establishing its accuracy and currency.[68]

Not surprisingly, therefore, in a science context, the lifecycle of a dataset, from acquisition to compilation and derivation, comprises important areas of concern to accuracy.[69] With respect to accuracy, *acquisition* is the most important stage in the lifecycle of a dataset, since it is the point where the original observations are collected and where "fundamental assumptions, calibrations and corrections are made."[70] *Compilation* is the stage where a database is created; it occurs when the data are assembled into some sort of comprehensive arrangement or into a scientific dataset, and it is a phase during which many errors can be introduced. *Derivation* is the stage where data are being manipulated; the output of this process is a representation, interpolations, averaging and any number of manipulative techniques that may change the form, format or structure of the data. This may or may not be a reversible phase and is a diversion point from the original observations. For this reason, keeping the raw data as well as derived data is important.

This is why, in a scientific context, data accuracy is critical and why the data need to be reliable. Data quality is normally articulated in a dataset's metadata; without metadata or data quality parameters, a scientist will not use, trust or rely on those data. Metadata are essential for the dissemination of scientific data whereby "a data set without metadata, or with metadata that do not support effective access and assessment of data lineage and quality, has little long-term

(ICDAT2002), December 19-20, 2002, Academia Sinica, Taipei, Taiwan, 53–64. Available at http://perpos.gtri.gatech.edu/publications/ICDAT2002.pdf#page=1. In this study, the author proposes a formal method for analyzing records management and archival procedures and systems to determine whether they maintain and preserve authentic records over time. The analysis procedure is based on a formalization of archival and diplomatic concepts and principles as definitions and axioms. Concepts such as digital record, record series and archival integrity are defined and axioms characterizing authentic documents and authentic records are formulated. A procedure is described for storing and retrieving the digital records of a record creator that incorporates elements to ensure the integrity and authenticity of the records. The theories of record integrity and authenticity are used with theories of communications security and belief to prove that the procedure achieves its goal of preserving the integrity and authenticity of the digital records.

[67] Guptill and Morrison, *Elements of Data Quality*, op. cit.

[68] Significant research describing the provenance of data in molecular genetics databases is one example of the importance of this concept for validating research (see Mark Greenwood et al. (2003), "Provenance of e-Science Experiments: Experience from Bioinformatics," in *Proceedings UK e-Science All Hands Meeting 2003*, Simon J. Cox, ed. Available at http://www.nesc.ac.uk/events/ahm2003/AHMCD/pdf/047.pdf).

[69] See Derek G. Clarke and David M. Clark, "Chapter 2: Lineage," in Guptill and Morrison, *Elements of Data Quality*, op. cit., 13–30.

[70] Ibid., 18.

use."[71] In fact, as the general study 10 findings demonstrate, data portal discovery services rely on metadata descriptions, which are seen as a form of "truth in labelling." For users of these portals, and indeed among scientists in general, it is considered "axiomatic that a database has limited utility unless the auxiliary information required to understand and use it correctly—the metadata—is included in the record."[72] In the sciences, metadata are also a means of attesting to and assessing a dataset's authenticity. In other words, authenticity in the sciences is linked to a clear lineage recorded in the accumulating metadata surrounding data, which closely parallels the situation with respect to digital records in general. Both data and their cumulative and related metadata must be present, clear, unambiguous and un-compromised. In the absence of metadata, it is possible to gain some understanding of a scientific dataset if there are associated peer review papers and reports that describe them; however, this would be a more laborious process.

As expected, one can find ample evidence of a concern for data accuracy and authentication in the legal arena. For example, satellite images have been admitted into evidence in a few legal cases. "The admissibility of remote sensing information must be examined within the context of the general requirements for admission of scientific evidence and expert opinion."[73] A litigator seeking the admission of remote sensing data as evidence must (1) qualify an expert, (2) authenticate and prove the contents of the data and (3) establish that proper and accepted processing techniques were employed. The use of an archive history file accompanying the final satellite imagery exhibit provides the potential for objective, external authentication and establishes that appropriate techniques and methodologies were employed in the creation of the exhibit. An archive history file is a document listing (1) all the data used in the creation of the final exhibit, (2) all the tools used in the creation of the final exhibit and (3) all the processes and methods used to create an exhibit.[74]

When one moves away from pure scientific data and into the realm of records as defined by archival theory, a more explicit concern for authenticity does surface. A good example here is the laboratory notebook. The Chemical Sciences Roundtable, among many other groups, has discussed the problems of moving from bound paper notebooks to electronic notebooks. Since these lab notebooks can be used to determine precedence in such things as patent cases, they are legal documents and their authenticity must be established to give them evidential value. As Roundtable panellists explained, "One of the purposes of an electronic notebook is to have a historical record that is used, among other things, for establishing priority and for integrity concerns in science."[75]

The concept of reliability in the sciences is also influenced by the focus on data rather than on records. Because of the focus on the *accuracy* of data, the concept of reliability is more likely to be used in reference to *collections* of data. Scientists presume scientific data to be reliable because they were collected by a federal or state agency or because of the professional reputation of the scientist who collected the data. In the sciences, the concept of reliability is closely associated with the concepts of reproducibility and accuracy. More generally, reliability is a quality that can be attributed to a person, as in a reliable person; to a device, such as a reliable

[71] National Research Council, *Preserving Scientific Data*, op. cit., 36.
[72] Ibid., 31, as cited in Lauriault and Craig, "General Study 10 Final Report," op. cit., 76.
[73] Sharon. H. Hodge (1997), "Satellite Data and Environmental Law: Technology Ripe for Litigation Application," *Pace Environmental Law Review* 14: 714. Hodge's article references the following cases: *United States v. Reserve Mining Co.*, 380 F.Supp. 11 (D.Minn. 1974) and *Gasser v. United States*, 14 Cl.Ct.476 (1988).
[74] Ibid.
[75] National Research Council, Chemical Sciences Roundtable, *Impact of Advances in Computing and Communications Technologies on Chemical Science and Technology: Report of a Workshop* (Washington, D.C.: National Academy Press, 1999), 173. Available at http://www.nap.edu/catalog.php?record_id=9591#toc.

machine; or to a system that is organized to accomplish certain ends, as in a reliable computer or records system. It is the individual assessor who determines what attributes are required before reliability can be reasonably inferred. Thus, to scientists, reliable data are data collected by a competent scientist using procedures and instruments that are reliable. Reliability is a matter of degree, however. The reliability of data is determined by examining information about their provenance, asking: Were the data created by a competent person; that is, a person who has professional credentials or is certified by a standards organization?

Thus, a set of data that was generated by a reliable person using trustworthy methodology and that remains complete and uncorrupted might be said to be "reliable." That would mean that the majority of its data points are correct and precise (accurate) and that their integrity has not been compromised. If the methodology was not sound, then the accuracy of the data may be low and the dataset would be unreliable. If the data are internally inconsistent—for example, by giving a wide range of data points or conflicting answers—then the data would be regarded as imprecise and thus unreliable. If the dataset is not complete or is otherwise corrupt, because of a failure at the point of collection to capture the full range of data or because of data loss at a later time, then the set is unreliable.

This difference in usage between the individual data point and the collective dataset is presumably a result of the scientific need for large collections of data. A single data point, although it may in fact be accurate, cannot be trusted by scientists to stand for a fact. A large set of data is needed for data to be reliable. The reason for this is that scientists do not expect absolute accuracy. Regardless of the precision of instruments and the soundness of the methodology, instruments can deliver noisy data leading to mistaken conclusions by the scientist. Therefore, single data points or small datasets cannot be trusted. Datasets need to be robust to be reliable so that the inevitable errors are diluted. Robustness implies large collections of data where the individual entries of data are complete.

In some situations, robustness can be a substitute for accuracy in producing a reliable dataset. Some measurements, such as certain measurements generated in the medical field, are difficult to make accurately. As one article notes, "The very nature of biomedical objects is volatile and irregular . . ."[76] In other areas of inquiry, such as meteorology, obtaining data quickly is more important than achieving high levels of precision and accuracy. When data are inaccurate, having numerous versions of the same observations can help smooth the inaccuracies. With a large dataset, it is also possible to see the overall precision of the data. Outliers can be seen for what they are and examined, thereby increasing the trustworthiness of the dataset. "Especially in areas of high data density, inaccuracies can be detected by humans or by computers from comparison with other data points, making it possible to bypass the inaccuracy."[77]

Authenticity, accuracy and reliability in the scientific sector case and general studies

The case and general studies carried out in Focus 2 tend to confirm the findings on authenticity, accuracy and reliability evident in the literature review. For example, case study 14

[76] A. Minitski, A. Mogilner, C. MacKnight and K. Rockwood (2003), "Data Integration and Knowledge Discovery in Biomedical Databases. Reliable Information from Unreliable Sources," *Data Science Journal* 2: 25. Available at http://journals.eecs.qub.ac.uk/codata/journal/contents/2_03/2_03pdfs/DS131.pdf.

[77] National Research Council, Steering Committee for the Study on the Long-term Retention of Selected Scientific and Technical Records of the Federal Government. *Study on the Long-term Retention of Selected Scientific and Technical Records of the Federal Government: Working Paper*s (Washington, D.C.: National Academy Press, 1995), 58. Available at http://books.nap.edu/catalog.php?record_id=9478#toc.

(Archaeological Records in a Geographical Information System) found: "There is more concern over the reliability and accuracy of the records than the authenticity."[78] It might be argued that authenticity is a by-product of this emphasis on accuracy, but, as a distinct concept, authenticity is not developed in the studies. The lack of a strong understanding of authenticity is a result of the scientific emphasis on data and information over the records that contain them.

In case study 14, the archaeologists involved in the operations of the GIS have only thought about the concepts of reliability, accuracy and authenticity in terms of the data rather than in terms of the record. From the creator's viewpoint, reliability and accuracy relate to the reliability of the data source, so it is assumed that if the source is reliable the data will be reliable and accurate. In fact, the archaeologists in this case study assumed authenticity on the grounds that the datasets were obtained from a state repository or from a researcher trusted as a professional who maintains information securely. Yet, there was a sense of discomfort with the concept of authenticity as applied to records because the archaeologists saw their work as an ongoing compilation that could not be broken up into discrete units and treated as records.

Case study 19 (Preservation and Authentication of Electronic Engineering and Manufacturing Records) reported on an engineering experiment to develop an open-source preservation format for digital computer-aided design (CAD) records of solid models used in high-tolerance manufacturing of complex assemblies. The business owners in this study use CAD records in the science-based manufacturing of high-assurance, high tolerance machined piece parts for the U.S. government. In their words, "there is a critical, unsolved business requirement to maintain authentic records over time to enable the production of the pieces as long as the business requires them, with the assurance that they meet the same strict standards (tolerances) as the original piece."[79] The intent of the experiment was to preserve not only the geometric specifications of the model but also its semantically encoded metadata, joined to make a "new logical preservation format" for archival purposes. By "logical preservation format," the experiment partners in this study meant a format encompassing not only the fixed form and content of information representing the model but also instructions encoded within its metadata so that reasoning engines of the future can conduct "proofs" against the object to authenticate it as fit to support the procedural action for which it was designed to be used. Because the digital objects are held by trusted parties in a secure environment and the overriding need is the ability to preserve the CAD records for use in manufacturing pieces accurately, authenticity in the archival sense of the concept was seen to be of less concern to the partners involved in this study.

As previously stated, this lack of a strong concept of authenticity does not mean that scientists are not concerned with the issue, merely that they express the concern in terms of accuracy, reliability or integrity. Case study 08, which looked at Mars Global Surveyor Data Records in the Planetary Data System said, "Project team members, PDS managers and engineers and other Planetary Scientists do not traditionally use the term authentic to characterize the data products that they create, maintain and use. They are concerned that the data records are complete, reliable, accurate, and that the integrity of the data record is

[78] Richard Pearce-Moses, Erin O'Meara and Randy Preston (2004), "InterPARES 2 Project - Case Study 14 Final Report: Archaeological Records in a Geographical Information System: Research in the American Southwest," 29. Available at http://www.interpares.org/display_file.cfm?doc=ip2_c14_final_report.pdf.

[79] Kenneth Hawkins (2006), "InterPARES 2 Project - Case Study 19 Final Report: Preservation and Authentication of Electronic Engineering and Manufacturing Records," 4. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs19_final_report.pdf.

assured."[80] To this end, there are data processing plans, manuals, specifications and workbooks to guide processing, transferring and data preparation. Further, the data are peer reviewed for accuracy and reliability and are validated through a system that also conducts checksums.

Integrity is particularly important in cases where data have passed through several processes after first being received. Metadata that support integrity are sometimes referred to here as lineage data; that is, a record of the stages through which the data have passed. Case study 06, Cybercartographic Atlas of Antarctica (CAA), is instructive on this point. To ensure reliability, authenticity and accuracy of the digital entities and documentation in the CAA, data are acquired from authoritative sources and are peer-reviewed. Each would have been assessed against the Elements of Spatial Data Quality, which include:

- lineage
- positional accuracy
- attribute/thematic accuracy
- completeness
- logical consistency
- semantic accuracy
- temporal information[81]

Authenticity in geography is captured in standard metadata as data lineage. Lineage, a mandatory metadata element, includes the history of a geographical dataset. Key elements in the metadata identify characteristics such as scale, accuracy, age and limitations on use. Within the geomatics profession, certain data management practices have also been adopted (e.g., inclusion of source data, documentation of source data rendered and how these data have been modified). The reputation of the institution or scientist is also a factor; thus, the CAA relies on the professional practices and authority of the institutions from which data are derived, and adheres to cartographic professional practices to choose the right level of data accuracy and to select cartographers for the right representation, a process that is very much reliant on metadata and professional practices. An editorial group reviews the content of the CAA to ensure thematic accuracy. In addition, the CAA production environment is protected by security measures such as physical security, password protection and careful control of access depending on type of user.

Case study 26 (MOST Satellite Mission) reiterates the concept that robustness in a dataset can replace reliability. Raw data received from the satellite will sometimes be corrupted by technical failures. However, by processing large collections of data, the scientists are able to deal with this problem. During the processing, the presence of the errors is diluted and the end results, called "reductions," are not affected. "Whether or not these false or corrupt data are included in the calculations for the reduction, does not affect the reliability of the outcome."[82]

The MOST scientists consider data that they receive "authentic" if there is no indication that the data received differ from the data recorded by the satellite instruments. The data are only accurate, however, to the extent that they truly represent the physical phenomena being observed, within the capability of the instruments. The researchers generate other data algorithmically from the original data and consider the derived data accurate to the degree that the algorithm transforms all the original data as expected.

[80] William Underwood (2005), "InterPARES 2 Project - Case Study 08 Final Report: Mars Global Surveyor Data Records in the Planetary Data System," 22. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs08_final_report.pdf.
[81] Lauriault and Hackett, "Case Study 06 Final Report," op. cit., 19.
[82] Bart Ballaux (2005), "InterPARES 2 Project - Case Study 26 Final Report: MOST Satellite Mission: Preservation of Space Telescope Data," 13. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs26_final_report.pdf.

General study 10 (Preservation Practices of Scientific Data Portals) was undertaken to collect information about the actual practices, standards and protocols currently used by broadly defined existing data services, archives, repositories, portals or catalogues in the sciences. Although the sample size from each scientific discipline is small, thus limiting cross-disciplinary analysis, the study does provide a deeper understanding of practices in the natural and physical sciences as these pertain to portals, selected case studies and their associated data, as well as an exploratory review that considers the importance of issues such as accuracy, reliability and authenticity in the management of scientific data exchanged through portals. The findings of general study 10 provide further evidence of the appreciation among scientists of the concept of authenticity, despite their general lack of familiarity with the term as it is used in an archival context.[83] Among the thirty-two scientific data portals surveyed in this study, the term "authentication" is often used, and many of the qualities of authenticity (as related to the concept in an archival context) are discussed despite the fact that the term "authenticity" is never used.

Observations derived from these few case studies suggest that accuracy is associated with the risk of having inaccurate data: the more legal requirements there are, the more rigorous are the quality checks. Also, the more automated the process is, the more technical the checksums are and the more reliant the creators are on the technical systems in place and the less reliant on human checks: this is the case with the NASA Mars Surveyor Data, the Engineering Drawing study and the MOST satellite data. Professional practice, however, is very important in the Cybercartographic and the Archaeology case studies, as is a reliance on the trust associated with the integrity and authority of external data providers.

Conclusions

There is no question that the concepts of authenticity, accuracy and reliability are important in the preservation of scientific records of science. Since sound scientific research is dependent on the accuracy of data gleaned from scientific experiments, it is logical that the concept of accuracy looms larger than the concepts of authenticity and reliability. Questions about the accuracy of the data maintained over time are not dissimilar to questions relating to the authenticity of records maintained over time. This is clearly evident in the cases where satellite data have been used as evidence in legal proceedings, as well as in the case of laboratory notebooks and in many of the scientific data portals surveyed in general study 10. Concerns for data lineage can be seen as analogous to archival concerns over provenance and the chain of custody, and the recognition that reliable datasets are connected to authoritative data collectors has echoes of archival concerns for the authority of records creators. The differences in scientific use of these three concepts is more one of emphasis reflecting the particular concerns of scientists, but there is no evidence of real disregard for the concepts of authenticity, accuracy and reliability as viewed from the archival perspective.

Another important finding is the relative importance that scientists place on the content of a record in terms of its data quality (i.e., the accuracy of its content) when appraising its long-term value, something that archivists have hitherto generally considered irrelevant when conducting appraisals. In fact, many scientists, especially those in geomatics, argue that the data quality of a record should be an important factor in the decision of what scientific data to preserve and that archivists must, therefore, consider data quality in their appraisals if they are to acquire data from

[83] Tracey P. Lauriault and Barbara L. Craig (2007), "InterPARES 2 Project - General Study 10 Final Report: Preservation Practices of Scientific Data Portals." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs10_final_report.zip.

the sciences. To this end, the fact that dataset users are expected to recognize that the analysis and interpretation of a dataset requires discipline-specific background knowledge and expertise suggests that archivists will fare better at archiving specific types of scientific data if they collaborate with scientists and specialists in the field. Alternatively, archivists can trust that either the scientists or the bodies managing the data will have already appraised the data in their custody; in which case, the archivists can instead work with scientists and their related institutions to add specific archiving practices into the data creation, management and preservation processes.

Compounding things further is the often disparate ways that the term "record" is understood and used by archivists and scientists. For many scientists, record is synonymous with data, databases and related information; entities that in an archival context are not generally considered records, except in very special and limited circumstances. As the Focus 2 research suggests, this is not, to a large degree, simply a matter of semantics; rather, it is a fundamental difference in perspective between scientists (creators) and archivists (preservers), exacerbated by the emergence in all disciplines of often highly ephemeral interactive and/or dynamic information that exists only in digital form. More importantly, it appears that the nature of the "record" within the digital environment may be changing dramatically. If so, traditional archival science will have to adapt to these changes in both theoretical and practical terms if it is to preserve this new information environment in the archives of the twenty first century.

In summary, although scientists indeed do recognize the importance of maintaining the "records" of scientific work in authentic, accurate and reliable form, it is important for both communities to be sensitive to differences in terminology usage as well as to fundamental conceptual differences regarding the very essence of what constitutes a record in both disciplines so they can work together to meet a common interest in long-term preservation.

Focus 1 – the Arts

Scope of the research

Although there is no lack of interesting questions about digital preservation in the sciences and government, the artistic creative activities contemplated by Focus 1 are so multifarious as to call into question some fundamental assumptions upon which the InterPARES Project was founded. Could such diverse digital objects be compared at all? Do artists working in what were historically different media share any common conceptions about preservation issues? Could the qualities of records identified by archival science have any consistent meaning to such diverse creators?

To come to grips with the theoretical and historical aspects of these issues, Domain 2 researchers combed the disciplinary literature for citations and explanations of concepts related to the identity, nature and preservation of artworks. Annotated bibliographies for music, dance, photography, moving images, sound recording, visual art, electronic literature, theatre and architecture were compiled, posted on the Project's Web site[84] and incorporated into the bibliographic database. Other bibliographic references were gathered from the case study reports.

Owing to certain weaknesses in the way the bibliographies were constructed, they served as research tools, not as final products of the Domain. InterPARES 2 did not have deep expertise in all disciplines, so some searches and annotations were more exhaustive and penetrating than

[84] See http://www.interpares.org/ip2/ip2_documents.cfm?cat=biblio.

others, and some bibliographies include items that are only tangentially relevant to the research questions of the Project. Also, some disciplines were quicker than others to adopt digital media and to realize the challenges that such a move posed for preservation, so many references deal with the concept of authenticity only in non-electronic media. Finally, since much of this work was conducted in the early stages of the InterPARES 2 Project (2002-2003), it does not reflect the most recent disciplinary thinking. To cite only a few examples, the bibliographies do not include major initiatives in electronic literature,[85] some significant developments in the philosophy of aesthetics[86] or descriptions of important preservation projects by the Variable Media Network and MUSTICA investigators in 2004. This report attempts to incorporate insights from such recent research.

Nevertheless, the bibliographies provided a good starting point for an historically informed analysis of the concepts. This was conducted by Domain 2 scholars of the arts and presented in several research papers. Their analysis is summarized below.

Grounding this conceptual survey was the work of the case studies researchers who investigated how selected creators thought about issues of identity, integrity and preservation in the context of specific works of art. Seven case studies treated the works of eleven creators, covering music, dance, theatre, moving images, interactive media installation and online publication. Although these disciplines have historically been fairly distinct, the Domain 2 researchers found that digital technology has fostered much interdisciplinary collaboration. The sharing of technology among disciplines has also helped reconcile different disciplinary conceptions about the nature of art and what needs to be preserved.

Conceptual analysis: authenticity, accuracy and reliability in the literature of the arts

Art practice and theory rarely concern themselves with by-products. To be sure, artists create (then often neglect) documents pertaining to their creative activities. Those documents that represent the transactional relations of artists and their patrons, such as commissions, contracts and correspondence, do not differ substantively from similarly functional documents of any other creator to the extent that their form and content are governed by the legal contexts in which they arise. But artists also generate other by-products that have no equivalent in the records of business, government or science. For example, most artists make and keep sketches—collections of ideas (sometimes fragmentary but sometimes even apparently complete works)—that are never published as final products. From an artistic sketch, unlike from a draft of a legal record, one might not determine the form of the work(s) that will result. This is partly because many sketches are merely components, not yet integrated into a whole. But the deeper reason is that artworks themselves are so different from records. They may have little to do with facts. They are complete and effective simply when the creator finishes them, not because they instantiate a fixed form or result from following a fixed procedure. Many are ephemeral, constituted essentially as experiences rather than as concrete documents or objects. And many are interactive, with their content and form determined partly by input from agents outside the artist's control.

[85] See, for example, Nick Montfort and Noah Wardrip-Fruin (2004), "Acid-Free Bits: Recommendations for Long-Lasting Electronic Literature," Version 1.0, June 14, 2004. The Electronic Literature Organization. Available at http://www.eliterature.org/pad/afb.html; and Alan Liu et al. (2005), "Born-Again Bits: A Framework for Migrating Electronic Literature," Version 1.1, August 5, 2005. The Electronic Literature Organization. Available at http://www.eliterature.org/pad/bab.html.

[86] Davies, *Art as Performance*, op. cit.

Not surprisingly, then, it is difficult for artists to relate archival conceptions of a record to their documents, digital or otherwise. They apply the terms "authenticity" and "accuracy" to final products instead—to the objects or experiences that are the focus of aesthetic appreciation. Certainly, it is important to know that an artwork is what it purports to be, and has not been forged, tampered with or otherwise corrupted; the identity and integrity of artworks are important historically and culturally and may affect their financial value.

For most artists and audiences of art, the word "authentic" carries a primary sense of "original." In diplomatics, an original is a record that is primitive,[87] complete and effective. However what constitutes an original artwork varies considerably with the diverse conceptions of the nature of art.

The notion of authenticity as originality is most straightforward, and conforms most clearly to diplomatic conceptions, for a "singular" artwork[88] that is a (relatively durable) physical object. Artists may date and sign these objects, and some artists include elements (a special symbol, or a self-portrait in a crowd scene) that brand the work with their identity. Some of these elements can be understood as intended to provide the work with authenticity by establishing its origin with the artist. However, since most such elements are easy to forge, the originality of a singular artwork is best established "by a complete and dependable record"[89] of where the object has been since it left the artist's hands. This record may also include information about alterations or "restorations" made to the object, thus addressing its integrity as well as its identity. In contrast to administrative records, whose originality can be determined by observing whether they manifest all the necessary elements of the documentary form that defines them, artworks need not conform to a pre-established form. Thus, provenance is the principal testament of originality; that is, of authenticity. In this connection, some of the artistic literature also uses the term "authentic" interchangeably with "genuine." "Genuineness is based on and reflects a direct causal relation to the artist."[90]

If no such record of provenance exists—as is frequently the case—the originality of an object can be judged by "expert . . . comparison with works already accepted and works already rejected as . . . by the same artist."[91] In effect, this judgment is an "authentication" of a work, like the authentication of a record—the declaration of its authenticity at a specific point in time by a juridical person entrusted with the authority to make such declaration—that is necessary when its authenticity cannot otherwise be presumed. An expert's examination of the materials of the work is analogous to a diplomatic analysis of the extrinsic elements of a document's form. Expert authentication may also consider the structure of the work—the relations of the parts to each other and their function in the whole—which is analogous to analyzing the intrinsic documentary form of records. But artworks, even in the same genre by the same artist, may differ widely from each other in these respects, and the many examples of expert-deceiving forgery teach us that the criteria for authentication based on intellectual structure are, at best, provisional.

---

[87] That is, the first complete and effective instantiation of the record.

[88] "Singular artworks are unique, occurring at only one place at a time. Paintings, collages, carved sculptures, and Polaroids are typical examples of singular works" (Guy Rohrbaugh, "Ontology of Art," in *The Routledge Companion to Aesthetics.* 2nd ed. B. Gaut and D.M. Lopes, eds. (New York: Routledge, 2005), 242. Online reprint available at http://web.mac.com/rohrbaugh/iWeb/Site/Philosophy_files/encyclopedia3.pdf).

[89] Nelson Goodman (1996), "Authenticity," in *The Dictionary of Art*, Jane Turner, ed. (New York: Grove, 1996), 834.

[90] Jerrold Levinson (1990), "Autographic and Allographic Art Revisited," in *Music, Art, and Metaphysics: Essays in Philosophical Aesthetics* (Ithaca: Cornell University Press, 1990), 106.

[91] Goodman, "Authenticity," op. cit.

Conceptions of authenticity become more complicated for "multiple" artworks, which can occur in different places at the same time; for example, literature, prints, music, films, dances and installation art. The authoritativeness and effectiveness of any occurrence of such a work do not depend on its "primitiveness," but the words "original" and "authentic" are still used to refer to the link between it and its creator. For example, to discredit unauthorized circulation of his novel, *Ulysses*, James Joyce provided the authorized publisher with a letter certifying that a particular edition—an unspecified number of physical objects—would be "the only authentic one."[92] For sound recordings that are cloned for publication, the term "master" is used instead of "original" to designate the authoritative source. Here "authenticity" is also used simply to indicate how exactly the copies reproduce the aural experience of the original recording, without regard to the original's status as a record.

Many of these multiple artworks result from executing instructions with specific instruments. The instructions can be executed with the instruments at different places or times, producing multiple objects or experiences that nevertheless all arise from the same procedure. For example, oily ink is rolled over a grease-pencil drawing on a moistened limestone plate, which is then pressed against paper; with each pressing, an instance of the work—a lithographic print—results. Moreover, some types of multiple works (music, dance, theatre) are created for performance: they are temporal experiences resulting from the actions of "interpreters" who execute the instructions with the specific instruments but who are also allowed by convention to add other information. In such contexts, the word "authentic" is used to indicate the causal link of the instructions and instruments to the creator. For example, a study of the emulation of an interactive video artwork, *Erl King*, observes that "the original [computer] code was written by the artists and their collaborators, and was therefore deemed critical to the authenticity of the work."[93]

For performed artworks, however, it is important to distinguish the authenticity of the documents conveying the instructions from what is called the authenticity of the performance. The latter entails notions of accuracy: for example, a musical performance is "accurate" to the degree that it realizes all the instructions in the score. Note, however, that scores do not make explicit all information necessary for accurate performance; performers must also adhere to implicit conventions of "performance practice," specific to the composer's time and place, that may modify the meanings of the symbols on the score. Some theorists call such accurate performance "authentic"—"a performance that reproduces all that is constitutive of the work's individuality."[94] Since the accurate and conventional execution of instructions produces an authentic instance of a multiple work, the notion of authenticity is detached from the property of primitiveness that characterizes the archival conception of originality.

Performance authenticity may be regarded as a matter of degree: to the extent that a performance is accurate as described above, it is "authentic." This contrasts with the meaning in diplomatics, in which only reliability is a question of degree, while authenticity is an absolute. Indeed, one might question whether performance authenticity has anything to do with preservation at all, since performances are ephemeral, not fixed. At least one can say, however,

---

[92] Letter excerpt from James Joyce to Bennett Cerf (April 2, 1932), cited in Robert Spoo (1998), "Copyright Protectionism and Its Discontents: The Case of James Joyce's *Ulysses* in America," *Yale Law Journal* 108(3): 659.

[93] Caitlin Jones (2004), "Does Hardware Dictate Meaning? Three Variable Media Conservation Case Studies," *horizon^0* 18(2). Available at http://www.horizonzero.ca/textsite/ghost.php?is=18&file=6&tlang=0.

[94] Stephen Davies (1991), "The Ontology of Musical Works and the Authenticity of their Performance," *Nous* 25: 21–41. Reprinted in *Themes in the Philosophy of Music* (New York: Oxford University Press, 2003), 74.

that performance authenticity is only possible if instructions and instruments (and knowledge of interpretative conventions) are authentically preserved.

Two further, opposing senses of "authenticity" stem from two complementary purposes of art, evident in the following statement: photographs "are authentic to the extent that they do justice to the facts of reality, and they are authentic in quite another sense by expressing the qualities of human experience by any means suitable to that purpose."[95]

The first meaning is verisimilitude. Viewers of a photograph may expect that the more fully it accommodates the detail, tonal range and perspective that they would perceive in the real object—the more the image is seen as the real thing is seen—the more truthful it is.[96] In this context, then, authenticity signifies accuracy—the quality of a work that facilitates the viewer seeing the photographed subject as if seeing the actual subject. The verisimilitude of a work may also be attributed to the way the work was made; for example, the strictures on props, sets and camera technique promulgated by the Danish filmmakers' collective Dogme95, whose "supreme goal is to force the truth out of . . . characters and settings."[97] Such emphasis on procedure recalls archival conceptions of reliability, which a record possesses if it is capable of standing "for the fact it is about,"[98] because it is authored by a competent person, created according to a controlled procedure, and complete in its form.

This idea of accuracy (conformance to perception) differs both from accuracy of performance—how exactly instructions are executed—and from scientific notions of accuracy, neither of which refer to how a subject is perceived. Also, it is undercut by many artists' realization that any record involves innumerable subjective decisions. In the words of the photographer Richard Avedon, "A portrait is not a likeness. The moment an emotion or fact is transformed into a photograph it is no longer a fact but an opinion. There is no such thing as inaccuracy in a photograph. All photographs are accurate. None of them is the truth."[99] Similarly, theorists demonstrate "the unreliability of the photograph as a record, and how much a construction it is . . . The photograph fails as a fact."[100]

The opposing meaning of authenticity reflects the belief that the primary purpose of art is to represent subjective experiences, not facts, an attitude that understandably prevails in such non-representational arts as music. In this context, authenticity denotes the degree to which an artwork manifests the individuality and essence of its creator or of the culture in which it was created. The term is used by critics "to bestow integrity, or its lack, on a performer, such that an 'authentic' performer exhibits realism, lack of pretence, or the like."[101] It might be said in such cases that the artist is the artwork, unmediated by any records. The prevalence of this notion of authenticity explains why many artists do not concern themselves with explicitly marking the identity of their works; to them it is inconceivable that anyone else either could or would produce art like theirs. Anything an artist makes (or directs the making of) is authentic, by this definition.

---

[95] Rudolph Arnheim (1993), "The Two Authenticities of the Photographic Media," *Journal of Aesthetics and Art Criticism* 51(4): 537.
[96] Jerry L. Thompson (2002), "Truth and Photography," *Yale Review* 90(1): 25–53.
[97] Lars von Trier and Thomas Vinterberg (1995), "The Vow of Chastity." Available at http://www.dogme95.dk/the_vow/vow.html.
[98] From the definition for "reliability" from the InterPARES 2 Terminology Database. Available at http://www.interpares.org/ip2/ip2_terminology_db.cfm.
[99] Richard Avedon, Foreword to *In the American West 1979-1984* (New York: Harry N. Abrams, 1985). Available at http://www.richardavedon.com/#mi=1&pt=0&pi=11019&p=-1&at=-1.
[100] Aphrodite Désirée Navab (2003), Review of *Transforming Images: How Photography Complicates the Picture*, by Barbara E. Savedoff. *Journal of Aesthetic Education* 37(2): 114–121.
[101] Allan Moore (2002), "Authenticity as Authentication," *Popular Music* 21(2): 210. Available at http://journals.cambridge.org/article_S0261143002002131.

Such "personal" authenticity may work against verisimilitude,[102] as evidenced by Dogme95's prohibition on crediting a film's director.[103]

However, the notion of authorship, and its role in establishing authenticity, is problematized by many works that incorporate mass-produced materials or that combine original work with excerpts from other artworks. The digital representation of sound and visuals greatly facilitates such appropriation. The aesthetic effect of such works is attributable to multiple authors, some of whom may not have intended it, and conflicts over copyright inevitably follow. "Questions about display and preservation require an interpretation of exactly what constitutes the work and who is authorized to make decisions that will shape how it is received."[104]

On the basis of this review of the literature, it is apparent that those who wish to preserve authentic artworks should be aware of the web of overlapping and sometimes contradictory meanings around the concepts of authenticity, accuracy and reliability. The terms are applied more frequently to final products than to the by-products of artistic creation. Authenticity refers to different properties depending upon what kind of artwork is being referenced. It is sometimes equated with accuracy. Reliability is almost never mentioned, although it is implicit in the relatively few instances of art-making procedures.

Little of the literature deals more than superficially with the problems of preserving digital artworks. For example, the Domain 2 researchers found numerous discussions of how the verisimilitude of photography is compromised by editing techniques, but the literature generally fails to address the practical aspects of how to create and manage digital images as reliable records and preserve their authenticity over the long term. However, the distinction between singular and multiple artworks makes an interesting and suggestive parallel to the distinction between fixed, physical records and the kind of ephemeral displays of information that are constituted by the execution of instructions in a computer system.[105] To understand this parallel more fully, and to investigate how problems of identity and integrity can affect the possibility of preserving digital objects, it is necessary to review InterPARES 2's analysis of actual digital art.

Authenticity, accuracy and reliability in the artistic sector case and general studies

Domain 2 had an especially rich source of information from the case studies in the creative and performing arts. Completed studies are listed below, tagged with the code assigned to them by the Project.

- CS01 Arbo Cyber, théâtre (?)
- CS02 Performance Artist Stelarc
- CS03 *Horizon Zero*/Zero Horizon Online Magazine and Media Database
- CS09(1) Digital Moving Images: Altair4 di Roma. A Multimedia Archaeological Project: *The House of Julius Polybius*
- CS09(2) Digital Moving Images: National Film Board of Canada
- CS09(3) Digital Moving Images: Commercial Film Studio
- CS09(4) Digital Moving Images: WGBH Boston
- CS10 *The Danube Exodus*: Interactive Multimedia Piece

[102] Peter Kivy. *Authenticities: Philosophical Reflections on Musical Performance* (Ithaca and London: Cornell University Press, 1995).
[103] von Trier and Vinterberg, "The Vow of Chastity," op. cit., vow no. 10.
[104] Martha Buskirk. *The Contingent Object of Contemporary Art* (Cambridge, MA: MIT Press, 2005), 23.
[105] A similar idea, without reference to the arts, is suggested in Helen Heslop, Simon Davis and Andrew Wilson (2002), "An Approach to the Preservation of Digital Records," National Archives of Australia. Available at http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm2-888.pdf.

- CS13 *Obsessed Again…*
- CS15 *Waking Dream*

Various perspectives on each case are posted on the InterPARES 2 Web site, including: (1) the proposal; (2) a final report that answers the 23 questions of the research instrument; (3) a characterization of the creation of digital objects; (4) an activity model (for selected case studies only); and (5) a diplomatic analysis of the digital objects, attempting to identify the presence of records as those were defined by InterPARES 1. Although the case study characterizations, activity models and diplomatic analyses are especially pertinent to the research questions of the Project's other Domains, they also cast light on conceptions of authenticity, reliability and accuracy in interactive and dynamic systems of the creative and performing arts. So do the answers to some of the twenty-three questions, which the Domain 2 researchers helped design for this purpose.

In addition to the case studies, three general studies cast further light on the research questions. Two Web-based surveys solicited the comments of composers (general study 04) and of photographers (general study 07) on the nature of the digital objects they create, and the problems they have encountered with maintenance and preservation. Also, the MUSTICA project (general study 03), a collaboration of InterPARES researchers with researchers in French music-composition studios, yielded valuable insights from institutions that have made the maintenance of interactive digital music a high priority.[106]

Consistent with the results of the conceptual analysis, case studies researchers found that artists did not always distinguish between products and by-products (that is, between publications and records) when thinking about authenticity, reliability and accuracy. In case study 03 (*HorizonZero*) and case study 09(1) (Altair4 di Roma), the final products were posted to servers, or distributed on disks, whereby they lost their archival bond to the objects that were created in the course of producing them. Artists are more concerned about preserving the final products than the by-products, although they recognize the necessity of the latter to the former. Indeed, it seems sensible to consider how to preserve both, especially since they involve many of the same issues of technological context, identity and integrity.

In any case, most of the artists studied by InterPARES 2 understand authenticity, first and foremost, to denote the causal link between them and the products or by-products of their activities. For example, to Stelarc (case study 02), a work is authentic if its content is his; any original performance (by him) is authentic but re-creations of the same actions by others would not be.[107] Similarly, the creators of the documentary Web site studied in case study 01 (Arbo) believe that authenticity is guaranteed if the artists who worked on the video and sound recordings during the original performances are the same who adapt them for the site. "'Authenticity' is maintained by the artist's constant presence."[108]

[106] See Jennifer Douglas (2006), "InterPARES 2 Project - General Study 03 Final Report: Preserving Interactive Digital Music - The MUSTICA Initiative." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs03_final_report.pdf; John Roeder (2006), "Authenticity of Digital Music: Key Insights from Interviews in the MUSTICA Project," version 2. Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs03_summary_report_ROEDER_v2.pdf; and Bruno Bachimont et al. (2003), "Preserving Interactive Digital Music: A Report on the MUSTICA Research Initiative," in *Proceedings of the Third International Conference on WEB Delivering of Music (WEDELMUSIC 2003)*, 15-17 September 2003, Leeds, UK (Washington, D.C.: IEEE Computer Society Press, 2003). Available at http://polaris.gseis.ucla.edu/blanchette/papers/wedelmusic.pdf.
[107] Henry Daniel and Cara Payne (2004), "InterPARES 2 Project - Case Study 02 Final Report: Performance Artist Stelarc." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs02_final_report.pdf.
[108] Martine Cardin (2004), "InterPARES 2 Project - Case Study 01 Final Report: Arbo Cyber, théâter (?)," 28. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_report_english.pdf.

Generally, this sort of authenticity is assumed to be ensured by the creators' control over the creation and organization of their digital objects, and by their marking the identity of those objects with metadata. For example, in the interactive multimedia installation of case study 10 (*The Danube Exodus*), authenticity, quality and reliability are guaranteed by the authors being able to oversee and control the publication or finalization, and then "stamp" the work with credits and copyright statements. It lasts only as long as the artists exercise stewardship over the product.[109] In strongly market-driven projects with rapidly changing tools, such as the commercial animation studio of case study 09(3)[110] and the contract-multimedia production company of case study 09(1),[111] where the incentive to preserve is almost nil, identity and integrity are ensured by restricting access to the creation or alteration of digital objects and by marking the objects with version numbers.

In auteur-driven projects, whose creators are concerned with their personal legacy, authenticity means that any supposed instance of the work appears the same as the original, according to the judgment of the creator. This recalls the academic definitions of "authentic performance" as one that accurately presents all the work's essential features. The importance of accuracy to the creators in case study 09(2) (National Film Board of Canada) is indicated by the "nervous breakdowns" some animators are reported to have suffered when confronted with versions of their work that had been degraded by migration to lower-quality display systems.[112] Some creators understood the notion of "reliability" in the same sense. In case study 01, the artists understood it to mean how well a video recording represented their conception of the work, with no regard to how it was made. They presume their records to be "reliable" because they believe the records are impossible to fake or, at least, that no one would want to do so.[113]

Often, in the most controlled contexts, "authenticity" is conceived purely as the "usability" of digital objects; that is, whether the objects (by-products) will function as expected in the software that is used to generate the final product.[114] In other words, authenticity is conflated with a kind of reliability. In collaborative efforts, the term "reliability" was not used independently at all,[115] or it was understood as a (desirable) characteristic of the systems that are displaying the documents, not as a (desirable) characteristic of the documents themselves. A reliable system, in this sense, displays the same information the same way every time it is called up.[116]

---

109 Sally Hubbard (2006), "InterPARES 2 Project - Case Study 10 Final Report: *The Danube Exodus*," 7–8. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs10_final_report.pdf.

110 James Turner et al. (2004), "InterPARES 2 Project - Case Study 09(3) Final Report: Digital Moving Images - Commercial Film Studio." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_final_report.pdf.

111 Isabella Orefice (2004), "InterPARES 2 Project - Case Study 09(1) Final Report: Digital Moving Images - Altair4 di Roma, A Multimedia Archaeological Project: The House of Julius Polybius." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-1_final_report.pdf.

112 Andrew Rodger (2006), "InterPARES 2 Project - Case Study 09(2) Final Report: Digital Moving Images - National Film Board of Canada," 11. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-2_final_report.pdf.

113 Cardin, "Case Study 01 Final Report," op. cit., 41.

114 See Orefice, "Case Study 09(1) Final Report," op. cit., 5; and Turner et al., "Case Study 09(3) Final Report," op. cit., 9, 18.

115 See Brent Lee (2004), "InterPARES 2 Project - Case Study 03 Final Report: HorizonZero/Zero Horizon Online Magazine and Media Database." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs03_final_report.pdf; Rodger, "Case Study 09(2) Final Report," op. cit.; Mary Ide (2005), "InterPARES 2 Project - Case Study 09(4) Final Report: Digital Moving Images - WGBH Boston." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-4_final_report.pdf; Hubbard, "Case Study 10 Final Report," op. cit.; J. Scott Amort (2004), "InterPARES 2 Project - Case Study 13 Final Report: *Obsessed Again....*" Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs13_final_report.pdf; Sydney Fels and Seth Dalby (2004), "InterPARES 2 Project - Case Study 15 Final Report: *Waking Dream*." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs15_final_report.pdf; and Nadine Hafner, Janine Johnston, Tracey Krause and Keum Hee Yu (2006), "InterPARES 2 Project - Case Study 22 Final Report: Electronic Café International (ECI)." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs22_final_report_DRAFT.pdf.

116 Daniel and Payne, "Case Study 02 Final Report," op. cit., 4.

In a somewhat different sense, reliability is a concern in works for performance, such as is noted in case studies 13 (*Obsessed Again…*) and 15 (*Waking Dream*), whose creators collaborate with performers, revising the work's instructions until they communicate the intention completely, and revising the digital instruments until they are adequate and capable of supporting repeated performances. However, both works considered in these two case studies were revised for each new performance, resulting in different versions of the work, so the digital objects were never fixed enough to apply the term.

It is interesting that diplomatic analysis of even the most interactive and dynamic artworks found the presence of records, or near-records, in the creators' systems, even when the creators did not conceive of their objects in that way. This can be explained by the fact that many artists working digitally save their files in some organized fashion, maintaining enough of a file-organization system to create an archival bond among the digital by-products of their activities Also, the interactive and dynamic features of the works are grounded in every case on fixed instructions and instruments. The work (the focus of appreciation) is interactive, but the records are not. Generally, however, these will not be reliable in the future, in the sense that they could be used to re-perform/re-generate the work, because they are tied to specific technical platforms and standards that change rapidly.[117]

The main preservation challenge, as many studies noted, is to preserve the technological context of these documents, or to find new technological contexts in which equivalent experiences can be generated. The question that hovered over all the case studies, therefore, was how equivalency; that is, accuracy, could be judged in the creator's absence. Here, the MUSTICA researchers' experiences are especially relevant. They all identify the necessity of preserving the instructions for producing, sequencing and processing sounds, and usually the sounds themselves, and they assert that their community uses a common "bedrock" of sound-processing procedures that should be migrated to any new technology. They also agree that a recording of the sound patterns does not preserve the work. No recording is "exact" or "precise," because it cannot manifest all the essential features of the work, because it records mistakes in performance, and because it cannot present the balance of sounds the composer has conceived for a live presentation of the music. Nonetheless, the MUSTICA interviewees regard recordings as essential to preservation, as the only substitute for the composer's authority after he or she has died.[118]

Conclusions and relevance of this analysis outside of the artistic sector

The various creative and performing arts converge in digital media works that combine physical objects, text, audio and moving and still visuals, all interacting with performers and audience. Preserving such "multiple" works means preserving the ability to perform (display) them. Not all artists embrace this conception or accept the limitations that digital media impose. Many artists are not concerned with preservation at all. But for those who are, the challenge is clear: creators need to take effort to specify and preserve the identity and integrity of the instructions and instruments, including their functionality, interoperability and accuracy of content, across technological change.

[117] See, for example, Rodger, "Case Study 09(2) Final Report," op. cit., 16.

[118] The French researchers affiliated with the MUSTICA project later proposed a detailed typology of the musical works, investigated the suitability of various metadata standards and proposed some methods of preservation (see Xavier Sirven (2004), "Authenticité et accessibilité des archives électroniques - MUSTICA, Le cas de la création musicale numérique," Technical Report, Université Technologique de Compiègne. Available at http://polaris.gseis.ucla.edu/blanchette/papers/RapportSirven.pdf.

This requires an understanding of all objects and their relations, along with the interdependencies of authors, performers and technology. The modeling activities conducted by InterPARES 2 were helpful in exposing these. Considering the experience the Domain 2 researchers had in recreating one of these works,[119] however, it is important to caution that there can be many subtle aspects to interactive systems that are only manifest when the creator evaluates the re-creation.

Artists' understanding of authenticity varies widely, and is often conflated with concepts more closely allied to reliability or accuracy. Nevertheless, many of them take at least some actions to identify the digital components of their works. Archival notions of authenticity are consistent with preservation intents and actions of the artists studied by the Domain 2 researchers. But these notions need to be nuanced in light of the disciplinary conceptions of authenticity and accuracy exposed in the Domain 2 analysis, as much in science as in the arts.

Multiple artworks, even those involving paper instructions and physical instruments, provide a model for how to regard the ephemeral "records" displayed by other digital information systems. These displays are performances of fixed instructions, using the instruments of the computer hardware.[120] Thus, methods for ensuring authenticity and reliability of multiple artworks can stand as a model for how those qualities can be preserved in digital record systems outside of the arts.

Relevance of the Benchmark Requirements of InterPARES 1

The focus of InterPARES 1 on digital records in administrative and legal systems, noted in the introduction, directed the derivation and content of the benchmark requirements it proposed. Since these documents are created in the context of well-defined procedures and function like paper documents with well-defined documentary forms, the benchmark requirements reflected long-established ideas about the authenticity of paper records. They assumed that recorded actions can be classified into types and that a record with a characteristic documentary form is associated with each type of action. Moreover, since InterPARES 1 focused on how to assess and maintain the authenticity of digital records once they become inactive and are selected for permanent preservation, it did not investigate how to create reliable digital records and maintain their authenticity during their active and semi-active life. That was the subject of a previous study, the "UBC Project,"[121] which was a collaboration between UBC researchers and the U.S. Department of Defense that produced the DoD Standard 5015.2 for recordkeeping systems.[122]

There are some difficulties, then, in applying the InterPARES 1 results to interactive and dynamic digital documents created by individuals or small collaborative groups in the arts and sciences. Although many such documents could be called "inactive," not all are records, and few of them (in particular, none in the case studies) have been selected for long-term preservation by an archival institution. This would suggest that the findings of the UBC Project might be more applicable, but the documents studied by InterPARES 2 also differ from those in the systems regulated by the DoD Standard. For instance, it is not clear how to classify the actions signified by the digital entities that are created as components of artworks, since the actions are steps in a generative process that may vary considerably from work-to-work and artist-to-artist.

[119] See the section later in this report titled "A strategy for preventing technological obsolescence of an artistic work."
[120] Duranti and Thibodeau, "The Concept of Record," op. cit.
[121] See http://www.interpares.org/UBCProject/index.htm.
[122] Since the DoD Standard was adopted, it can be said that InterPARES-related guidelines have been validated by their regular use in some government recordkeeping activities.

Concomitantly, it is not clear whether such various entities have any consistent documentary form that could be examined to determine whether they did participate in the creation of an artwork. Lastly, the UBC Project did not contemplate the special problems of interactive and dynamic systems. Informed by InterPARES 2 research, Duranti and Thibodeau's rethinking of the concept of a record deals with many of these theoretical issues.[123]

Nevertheless, some evidence for the relevance of the InterPARES 1 findings can be seen in the fact that even creators furthest from recordkeeping bureaucracy show an awareness of authenticity requirements in the way that they create and organize their digital objects. For instance, benchmark requirement A.1 asserts that authenticity can be presumed if certain identifying attributes are explicitly expressed and inextricably linked to every record. For the digital entities analyzed diplomatically by InterPARES 2, many of these attributes are at least implied, and often standard, as in some of the scientific datasets.[124] Even when the attributes are not explicit, it seems like a small step to include them (for instance, an historical trace of provenance) as part of the objects' metadata.[125]

Consider, moreover, that in several of the InterPARES 2 case studies the creators attempted to maintain their documents and encountered various difficulties. To the extent that those difficulties can be attributed to violations of the requirements proposed by the UBC Project and InterPARES 1, the requirements can be understood as relevant. For example, the Domain 2 researchers observed that various problems of archival bond (one of the attributes of record identity required to be explicit and linked by benchmark requirement A.1) can beset the digital entities associated with Web sites. In some cases, such as HorizonZero, they are not set aside in a recordkeeping system with other records with which they could form an archival bond.[126] In others, such as on the Legacoop of Bologna's site, the "entities on the Web site do not possess an archival bond beyond a chronological record of their posting."[127] Without these bonds, a creator may be able to maintain a publication as a final product, but the traces of its creation will be obscure. In contrast, the data files representing transactions with the Irish Revenue On-Line Service are structured to form natural aggregations "wrapped" together by addressee.[128]

Dynamic systems that draw information from constantly changing sources naturally run afoul of the benchmark requirements. If their displays are to serve as records of actions, the data they display must be fixed, or at least bounded,[129] and dated to enable redisplay. For example, in cases such as VanMap, the lack of date-stamping can prohibit preservation.[130] Another problem is exemplified by Stelarc's Web site, which the artist intends as a record of his work. One of its pages involves an interactive interface that simulates a performance of his work "Ping Body"[131] by requesting the "ping" (response) time from a server in Australia[132] to a randomly selected

[123] Duranti and Thibodeau, "The Concept of Record," op. cit.
[124] See, for example, Underwood, "Case Study 08 Final Report," op. cit.; and Ballaux, "Case Study 26 Final Report," op. cit.
[125] For such a proposal for digital artworks, see Alena Williams, "Rhizome.org," in *Permanence Through Change: The Variable Media Approach*. Alan Depocas, Jon Ippolito, and Caitlin Jones, eds. (New York: Guggenheim Museum Publications, 2003), 39–41. Online reprint available at http://variablemedia.net/pdf/Permanence.pdf.
[126] Tracey Krause (2006), "InterPARES 2 Project - Case Study 03 Diplomatic Analysis: *HorizonZero*/Zero Horizon Online Magazine and Media Database." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs03_diplomatic_analysis.pdf.
[127] Carolyn Petrie (2006), "InterPARES 2 Project - Case Study 25 Diplomatic Analysis: Legacoop of Bologna Web Site," 4. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs25_diplomatic_analysis.pdf.
[128] Tracey Krause (2005), "InterPARES 2 Project - Case Study 20 Diplomatic Analysis: Revenue On-Line Service (ROS)," 3. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs20_diplomatic_analysis.pdf.
[129] See discussion of the concept of "bounded variability" in Duranti and Thibodeau, "The Concept of Record," op. cit., 47–48.
[130] McLellan, "Case Study 24 Final Report," op. cit., 31.
[131] See http://www.stelarc.va.com.au/pingbody/ping.html.
[132] http://www.merlin.com.au.

remote Web server. The returned value controls the motion of wire-frame body limbs displayed on the screen, simulating the actual performances, in which Stelarc's own limbs are controlled by electrical shocks proportionate to the ping values. However (at least in July 2006), the Australian server no longer responds to the request, so the interface no longer accurately simulates the performance.[133] Another case study that involves such dynamic documents is the Cybercartographic Atlas (case study 06), and the general study of scientific data portals (general study 10) reveals that many of them have analogous external dependencies.[134] None of the benchmark requirements directly addresses this situation. Stelarc's site itself has not undergone technical modification (benchmark requirement A.1). In some senses, the technological context has not changed (benchmark requirement A.4)—the Australian server still exists, and the simulation (a Shockwave movie) still runs. In another sense, however, it has been modified to the extent that the Australian server administrator has removed the software routines from which the simulation requests the ping values. This is a rather subtle change (indeed, Stelarc's Web site administrator has not noticed it) and demonstrates the need for careful analysis of the inputs to dynamic documents.

Of course, providing for changes in technological context, as demanded by benchmark requirement A.4, is the most pressing problem for preserving all sorts of digital systems. Most proposals for preservation in the literature deal principally with this issue, which is also considered in the following section of this report. In InterPARES 2's studies of artworks, such as *Obsessed Again...* (case study 13) and *Waking Dream* (case study 15), and for the musical works studied by MUSTICA (general study 03), the creators are not truly maintaining their original works, but are essentially creating new versions—that differ in essential ways from the originals—by rewriting software for the latest technologies. They have not specified their works in ways that minimize or eliminate dependence on custom, proprietary or obsolescent instruments. Analogously, in case study 19 (Preservation and Authentication of Electronic Engineering and Manufacturing Records) from the science focus, the methods that the creators experimented with to verify the identity and functionality of machine parts specified by CAD (computer-aided design) documents could not be successfully realized without dependence on a proprietary reasoning engine that could not itself be preserved.[135]

Thus, even though the concepts employed by InterPARES 1 and the UBC Project are not entirely adequate for the systems studied by InterPARES 2, the benchmark requirements seem relevant, because a failure to follow them prohibits preservation, and because efforts to preserve include some of the actions they specify. There are, however, indications that they are necessary, if perhaps not sufficient, so a more thorough review of this issue seems warranted.

Experience with a Possible Maintenance Strategy

Issues

In activity that produces records, the identity and integrity of the records are not in question as long as the creators are still actively referring to them, because records that the creator relies on in the usual and ordinary course of business are presumed authentic. In some cases, however, the digital objects that are the components of these records are set aside and left inactive long

---

[133] There are other outdated/nonfunctional links, as well, that detract from the integrity of the site.
[134] See Lauriault and Craig (2007), "General Study 10 Final Report," op. cit.
[135] Hawkins, "Case Study 19 Final Report," op. cit., 8.

enough that technological change renders them unusable. In effect, even if their actual file structure and content may have been physically preserved, their integrity is undercut by the disappearance of the technological context needed to display them.

Technological obsolescence, as remarked above, is one of the primary concerns for creators, users and preservers of digital records. When such objects are merely backed up (which some confuse with "archiving"), a change of technological context is not evident. The loss of integrity will only be evident to the extent that records keepers and preservers monitor the authenticity of records when they are transmitted across space or time. What procedures of creation and transmission would ensure that these records will continue to be recognized as authentic?

Some proposed strategies, and their relative advantages and disadvantages, are summarized concisely by Heslop, Davis and Wilson.[136] These authors' suggestion—to require all digital components to be expressed in public-domain formats—is a good one, but does not address the special problems of custom-formatted entities like the ones often encountered in artistic activities.[137] It would be futile to insist, for example, that artists restrict their means of expression to the lowest-common-denominator formats. And the authors do not consider how or if such a strategy could maintain the interactive and dynamic attributes of records, for which there are no standard representations.

A strategy for preventing technological obsolescence of an artistic work

To consider more fully the problems of preserving documents with these special attributes, Domain 2 researchers attempted to resurrect a work that had already fallen victim to technological obsolescence: Keith Hamel's *Obsessed Again…* for bassoon and interactive electronics (1992), the subject of case study 13. The instructions and instruments specified originally by the composer are represented schematically in Figure 1. It was assumed that the musical score (the instructions for the bassoon, in portable document format) can be preserved, and that an accurate and reliable bassoon, microphone and amplification system will exist in the future. A recording of a performance of the work (in a format with freely available specifications, so presumably preservable) was also available.[138] However, the other instruments shown in the centre right of the figure are now obsolete. The sounds that the computer causes to be played during a performance—including their timing and their interaction with the sounds that the bassoonist plays—are encoded in the instructions symbolized as "code" in the figure. But the interactions are nowhere explicit; they can only be deduced by analyzing the code and listening to the recording. To be realized, they require a functioning software environment (the proprietary MAX 2.0 running on a proprietary operating system) to interpret them. Outside of that specific technological context they are inoperative, and it is difficult to discern what they are supposed to do without an intimate knowledge of the technical specifications of the hardware and of the syntax and semantics of MAX.

---

136 Heslop et al., "An Approach to the Preservation of Digital Records," op. cit.

137 See, for example, Nicola Bernardini and Alvise Vidolin (2005), "Sustainable Live Electroacoustic Music," *eContact!* 8(3). Available at http://cec.concordia.ca/econtact/8_3/bernardini_vidolin.html; and Joel Chadabe (2001), "Preserving Performances of Electronic Music," *Journal of New Music Research* 30(4): 303–305.

138 For an analysis of the terminology used to characterize various levels of software "openness," see Evelyn Peters McLellan (2006), "InterPARES 2 Project - General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_english.pdf. French language version available at http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_french.pdf.

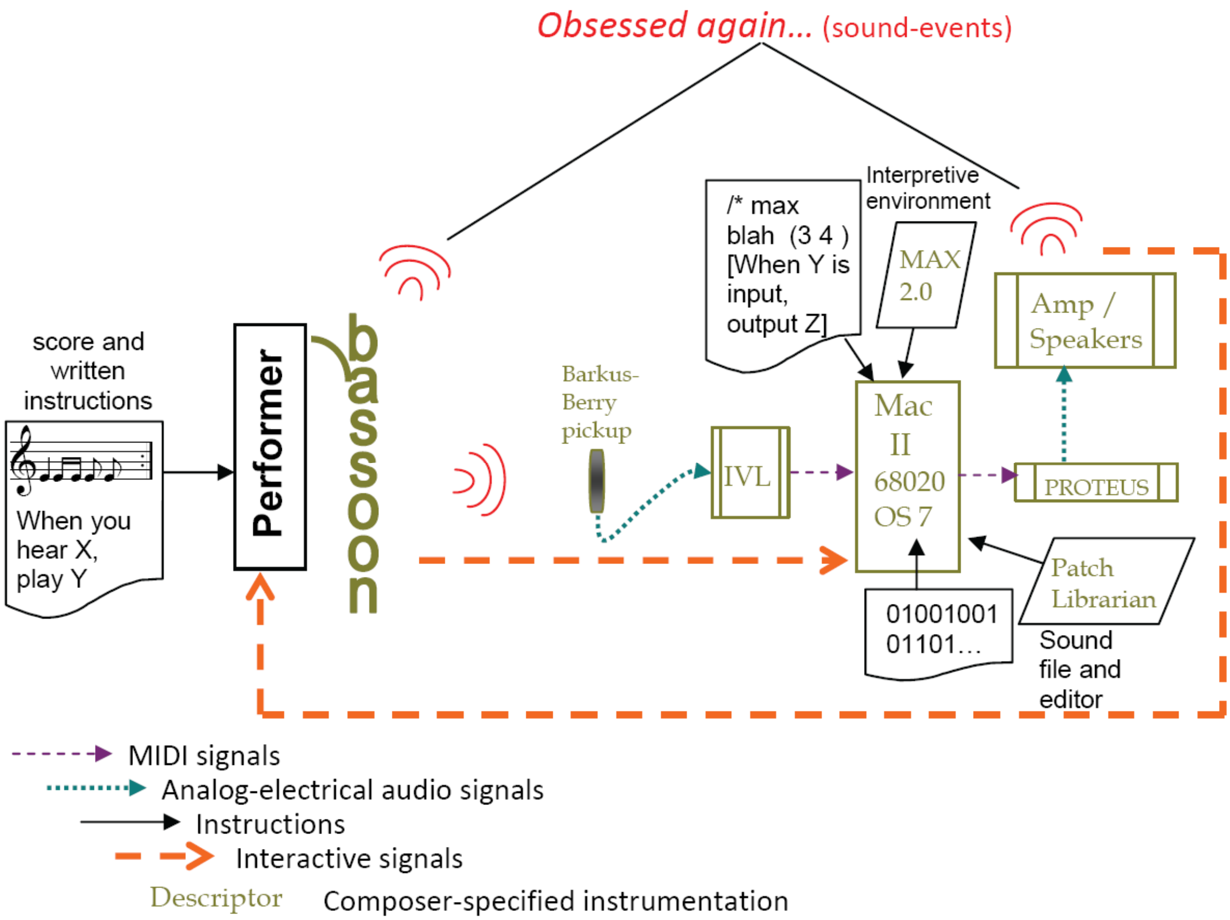


Figure 1. Schematic of Composer Instructions and Instrumentation Specifications for *Obsessed Again…*

The following paraphrased remark by one of the interviewees of the MUSTICA study summarizes the problem succinctly: 'The death of a patch [that is, the hardware instructions] means the death of the composition.' Ironically, at IRCAM, one of the institutions participating in MUSTICA, preservation efforts have produced meticulous documentation of how to perform works[139] but not enough information about content; thus, if hardware instructions no longer function, the integrity of the digital components is lost. Similar difficulties prevent the preservation of many interactive artworks. They have been articulated, for ephemeral art, by the Variable Media Initiative,[140] and the Electronic Literature Organization produced a substantial study of related issues in e-literature.[141] But the problems clearly extend to any system threatened by software or hardware obsolescence, including cases studied in the science focus (such as case study 19). A so-called "open" format can be proprietary and thus become obsolete if the proprietor ceases to support the format or asserts intellectual property rights that impede preservative transformations. Finally, even open, non-proprietary formats may become obsolete if future technology works differently than that of today.

[139] See, for example, Andrew Gerzso, "Performance Handbook: *Anthèmes 2* [by Pierre Boulez]," (Paris: IRCAM, 2005). Available at http://mustica.ircam.fr/mustica_1.2.0/rendu/pdf/output/Anthemes_2.pdf.
[140] See http://www.variablemedia.net.
[141] Liu et al., "Born-Again Bits," op. cit.

Resurrection of *Obsessed Again…* involved an exercise in controlled migration that simulated the transmission of records across space and time. One researcher, familiar with the technical details of the original instruments, translated the instructions from code into technologically neutral natural language stored in a word-processing document in non-proprietary format. Another researcher, armed only with these new instructions and the recording and with no other knowledge of the work or contact with the composer, wrote software that would control modern instruments to produce the same sounds and interactions that the original instructions and instruments did. Lastly, the Domain 2 researchers asked the composer, who was otherwise absent from the exercise, to judge the authenticity of a performance that employed the new instructions and instruments. By these means the researchers sought to establish a set of records and observe the effects of hardware and software evolution on them to determine whether it is possible to represent all that is essential to the work's identity in a technologically neutral way.

The responses of the composer confirmed the successes of this exercise while clarifying its limitations. He acknowledged that the machines and software in the new version interacted correctly with the bassoon's music, but he pointed out certain deficiencies that made the result somewhat different than he intended. That they affected authenticity was evident from his comment: "I like it, but it's not my piece." First, the sensitivity of the devices that receive input—in this case, the part of the system that detects the bassoon's sounds and translates them into digital inputs to the computer—were crucial to achieving the intended interactions. If the representation is too coarse-grained or fine-grained, the system may not respond when it should, or it may respond when it is not intended to. This sensitivity needs to be made explicit in the instructions for the artwork; it involves timing as well as other measurable aspects of the input. Also, the resolution of the output was crucial. In this case, the sounds that the electronic devices produced were all encoded in a proprietary format that could not be described in a technologically neutral way, the modern sound-producing devices could not be made to match those on the recording exactly, and the original instructions gave no indication of how accurately those sounds needed to be reproduced. Not surprisingly, the resulting sounds did not match the composer's intentions, and, in fact, this was the only reason he gave for not acknowledging the new performance as authentic.

For this work, it is not hard to imagine a solution. Recordings of the necessary sounds could be stored in a format with freely available specifications so that the migration would only involve reprogramming the interactions, which the researchers successfully did. And the very exercise confirms the intuition of the Variable Media researchers of how important it is to get the creator's feedback on attempted migration. In terms of the conceptual analysis above, it can be concluded that by making inauthentic performances one can discover how to provide instructions that can be preserved authentically and that can produce authentic performances. This supports the finding of Domain 3 that "preservation begins at creation."[142] Creators, while they are still living, are the best arbiters of the authenticity of performances. So it behooves them to describe their works in technologically independent (and authentically preservable) ways that will allow authentic performance in the future.

[142] Domain 3 Task Force Report, 164.

Analogies to a mechanical engineering case[143]

This experience has an interesting parallel in a very different case study, that of the Preservation and Authentication of Electronic Engineering and Manufacturing Records (case study 19). This case study, in fact, was also an experiment, initiated by the records creator to find a method of preserving active records to meet the creator's needs. In this case, the creator's need is to be able to use the records for the same purpose for which they were created: to manufacture piece parts for physical equipment. The equipment is often maintained for decades after its manufacture. At any point over this time it may be necessary to produce replacement parts if an existing part is damaged or wears out. The replacement part must fit into the piece of equipment exactly as the original part did. Piece parts are manufactured according to specifications produced as computer-assisted design (CAD) records using computer-assisted manufacturing (CAM) records that control the processes executed by robotic machine tools to manufacture parts with the right size, shape and configuration.

CAD/CAM systems today are proprietary and subject to obsolescence. The experiment was designed to test whether the CAD records could be translated from their proprietary format into a persistent format and preserved for use in some future, unknown CAM system to produce identical replacement parts. The formats chosen for preservation were independent of any specific hardware or software, freely available, standard and self-describing.

The experiment consisted of the records creator producing persistent format versions of its original records, transmitting them to a trusted digital repository as a surrogate for an archives, retrieving them from the repository and determining whether the preserved records could be used to produce the piece parts they described. The experimental design intentionally included several potential points of failure: the translation from propriety into persistent formats, transmission of the persistent records to the surrogate archives, ingest into these "archives," preservation, retrieval and return of the records to the creator and their use in production of replacement parts.

In fact, the experiment encountered failure at the first point. Even though the records creator employed two different types of freely available, standardized, self-describing formats to capture the piece parts, the persistent format records were not adequate to enable manufacture of replacement parts. This failure obviously entails inability to use the persistent format records to produce replacement piece parts. The intermediate steps in the experiment were executed without problems.

Unlike the *Obsessed Again…* case, there was no element of subjective judgment in the determination that the CAM experiment failed. However, there are parallels between the situation in the arts and that in engineering: the lack of an adequate language for expressing the specifications or instructions in the original records in a preservable format that could be used to perform or produce something that satisfied the original intent. Note that both failures were in specific cases. They do not amount to a failure of transformation to persistent formats as a preservation methodology. Rather, they identify specific areas where additional efforts are required.

Connections to the goals of the Project

Both of these experiments highlight the importance of the revisionary conceptual work that was the principal activity of all of the research domains in InterPARES 2. Considering the

[143] Kenneth Thibodeau contributed the content of this subsection.

interactive and dynamic environments exemplified by the various case studies, researchers were led to propose expansions to the traditional conceptions of record and metadata. Records in such environments as *Obsessed Again…* and case study 19 often encompass discrete components distributed across systems, while their behaviour, operations and even their authority to reside within computing environments may depend on the messages or instructions their metadata communicate to those same environments.

The experiments also confirmed the need, suggested by the conceptual analysis, for expansions to the traditional conceptions of authenticity, reliability and accuracy. The attempted resurrection of *Obsessed Again…* suggested that, in principle, a performance of a born-digital composition could be authentically and accurately performed sometime later using newly supplied elements (new samples, etc.). Case study 19 actually proposed an expansion of the underlying basis of presumed authenticity, saying that it depended not only on reliably populated attributes evidencing identity and integrity, but also on the conduct of "proofs" involving the semantic relationships of those attributes within a domain-specific ontology.

Toward Guidelines for Creating and Maintaining Authentic and Reliable Digital Records

Although Domain 2's bibliographic research found many theoretical discussions of the challenges posed to the authenticity and reliability of digital objects, it also found, as noted by the study of the digital recordkeeping practices of photographers who operate in artistic, scientific and governmental environments, that "documentation of procedures to create and preserve [records] in the digital environment for the long term has been sparse."[144] Given the urgency of the preservation problems identified in the introduction to this report, it seemed imperative that InterPARES 2 issue guidelines to assist creators in creating and maintaining preservable digital materials, especially records. It fell to Domain 2 to produce the *Creator Guidelines*[145] a document designed to accompany the *Principles for Records Creators*[146] developed by the Policy Cross-domain.

During development of the *Guidelines*, it became evident, from the conceptual analysis, case studies and general studies of the Project, as well as from the experiments described in the previous section, that certain principles would need to guide the content, form and presentation of the guidelines:

- They should reflect the concepts and practice of archival science; for example, distinguishing backups or repositories from archives.
- They should specifically address records of interactive and dynamic systems. For example, it is not sufficient simply to require documents to be in a format that is non-proprietary or that has freely available specifications, because no freely available description standard for interactivity yet exists.
- They should avoid using the terms authenticity and reliability, while still clarifying what the records must have to be authentic and reliable. This is because Domain 2 found that these terms, although precisely defined in archival science, mean different things to different creators, and that (if they are used at all) they are often confused or conflated.

[144] Bushey and Braun, "General Study 07 Final Report," op. cit., 3.
[145] See Appendix 20. The Guidelines also are available in booklet form at http://www.interpares.org/display_file.cfm?doc=ip2(pub)creator_guidelines_booklet.pdf.
[146] See the *Policy Framework* in Appendix 19.

- They should be worded so to make it clear (if not simple) what is required to satisfy them, even to such disparate creators as artists, scientists and bureaucrats.
- They should reflect the finding that, for records to be preserved, information and processes must be incorporated into their creation that will allow their identity and integrity to be ascertained in the future.
- They should be consistent, as far as possible, with guidelines issued by professional organizations, curatorial institutions and standards organizations.
- They should facilitate respect for cultural differences, freedom of expression, freedom of inquiry and right to privacy.

The *Guidelines* were worked out through an iterative method. On the basis of bibliographic research that exposed previous attempts at guidelines, candidate guidelines were proposed and considered by the InterPARES 2 International Team, considering the principles articulated above. The results reflect a consensus of archival scholars, practicing archivists and specialists in the arts, science and government focuses. Although it is presumed these guidelines apply to a large class of record-making and recordkeeping activities, the InterPARES researchers do not claim that the guidelines exhaust all of the preservation-related issues and concerns that may be associated with, or impacted by, records creation and maintenance activities. Thus, although the requirements for record-making and recordkeeping derived from them seem necessary, it cannot be claimed that they are sufficient for all cases; only experience will tell.

Other products of InterPARES 2 are also intended to assist in the creation, maintenance and long-term preservation of authentic digital records. The Metadata and Archival Description Registry and Analysis System (MADRAS) supports and eases the tasks of identifying, registering, describing and evaluating existing standards for the intellectual control of records from the moment of their creation throughout their appraisal and preservation.[147] InterPARES has also produced frameworks for the development of policies, strategies and standards regarding creation, maintenance and preservation of digital records; one framework is for organizations creating digital materials, and the other is for archival institutions or programs.[148] The Project's two models of records preservation—one reflecting a record lifecycle point of view (Chain of Preservation Model) and the other reflecting a record continuum point of view (Business-driven Recordkeeping Model)—can help organizations clarify needed procedures and resources.[149] Finally, the Terminology Database, which defines the terms used in the InterPARES Project, also includes a comparison with terms in existing dictionaries of all disciplines involved in the Project, thus fostering communication among creators and preservers of our digital legacy.[150]

---

[147] MADRAS is discussed at length in the Description Cross-domain Task Force Report. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_6_description_task_force.pdf.

[148] The context for this framework, known as the *Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records* (a.k.a., *Policy Framework*), is discussed in the Policy Cross-domain Task Force Report (available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_7_policy_task_force.pdf), while the framework itself is provided in Appendix 19.

[149] Narratives for both models are provided in the Modeling Cross-domain Task Force Report (available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_5_modeling_task_force.pdf), while the model diagrams and definitions can be found in Appendices 14 (COP Model, available at http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_14.pdf) and 15 (BDR Model, available at http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_15.pdf). Both models are also available on the InterPARES Web site at http://www.interpares.org/ip2/ip2_models.cfm.

[150] More detailed description about the Terminology Database and each of its components is provided in the Terminology Cross-domain Task Force Report (available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_8_terminology_task_force.pdf, while the Database itself is available on the InterPARES Web site at http://www.interpares.org/ip2/ip2_terminology_db.cfm.

PART FOUR

METHODS OF APPRAISAL AND PRESERVATION

Domain 3 Task Force Report

by

Yvette Hackett, Library and Archives Canada

Introduction[1]

Of the wide range of research areas that the second InterPARES Project broached, appraisal and preservation issues are considered by many to represent the core of the archival profession. Appraisal is that key function where the archivist's decision about what to acquire, and consequently, what not to acquire, establishes what primary records will be available in the future to support legal actions, historical research, genealogy—matters of identity, culture, history and rights.

Preservation decisions will either: ensure the successful survival of selected records over the long-term; preserve them in a manner that strips them of credibility and subsequent usefulness; or destroy them through error or omission, thus unintentionally overturning the appraisal decision. For many archivists, appraisal and preservation issues have the greatest impact on day-to-day work.

Background and mandate

Despite the importance of both appraisal and preservation to archivists, Domain 3 seemed to suffer more difficulties than did the other domains in agreeing on the focus of its activities and establishing a work plan. As set out in the initial research proposal, the Domain 3 research unit would begin by working with two products developed by InterPARES 1. It would first merge the Appraisal Task Force's *Model of the Selection Function*[2] and the Preservation Task Force's *Model of the Preservation Function*[3] with the *Manage Archival Fonds* model produced by the UBC-MAS Project.[4] The resultant "Grand Unified Model," as it was initially called, would represent the complete life-cycle of a record from its initial generation by the creator to its long-term preservation and access while in the trusted custody of the preserver. Once completed, prototyping experiments could then be conducted to illustrate how the functions depicted in the model could be incorporated into software applications.

Organizationally, there were difficulties integrating the highly experienced modelers from InterPARES 1, who were essentially continuing work begun in 1999, with InterPARES 2's new recruits, many of whom were unfamiliar with both archival activities and with the intricacies of the modeling methodology adopted by the Project. This situation was resolved in June 2003, when the modeling activity was moved into its own research unit (Modeling Cross-domain). The subsequent decision in February 2004 to create a new, second model based on the concept of the records continuum rather than on the lifecycle further distanced the work being done on the models from the researchers in Domain 3.[5]

---

[1] The author acknowledges the general contribution of all members of Domain 3 in the preparation of this report. In particular, the author thanks Luciana Duranti, Ken Thibodeau and Randy Preston for their contributions to the text and their editorial guidance. Any errors of representation or omission are the responsibility of the author.

[2] See Appraisal Task Force (2001), "Appendix 4: A Model of the Selection Function," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 239–252. Online reprint available at http://www.interpares.org/display_file.cfm?doc=ip1_aptf_model.pdf.

[3] See Preservation Task Force (2002), "Appendix 5: A Model of the Preservation Function," version 6.0, ibid., 253–292. Online reprint available at http://www.interpares.org/display_file.cfm?doc=ip1_ptf_model.pdf.

[4] See "Appendix B: Activity Models," in Luciana Duranti, Terry Eastwood and Heather MacNeil, *Preservation of the Integrity of Electronic Records* (Dordrecht: Kluwer Academic Publishers, 2002), 92–106. Online reprint available at http://www.interpares.org/UBCProject/a-0f.htm.

[5] See the Modeling Cross-domain Task Force Report. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_5_modeling_task_force.pdf.

The Domain 3 researchers also abandoned earlier plans to use InterPARES 1 case studies to validate the InterPARES 1 models prior to any system prototyping activity. The changes required in the original InterPARES 1 models to create the merged Chain of Preservation (COP) model meant the two original models were too out of date to provide useful input to the current work. In the case of the Business-driven Recordkeeping (BDR) model, no analysis could be undertaken until it was completed, which, given its late start, could only occur late in the Project.

Furthermore, as InterPARES 2 case studies were approved, it became obvious that InterPARES 1 case studies could not be used as a shortcut to get started on Domain 3 work while new case studies were being completed. The record-making environments being proposed for study were radically different from anything studied in InterPARES 1, even in the government focus. As with the modeling group, only a few members of Domain 3 had participated in InterPARES 1 and had extensive knowledge of the earlier case studies. It appeared that the learning curve would be daunting, while producing little insight relevant to the dynamic, interactive and experiential systems under study in InterPARES 2.

By the time the Midterm Progress Report was produced in the spring of 2004,[6] Domain 3 had little to nothing left of its original statement of work:

> ...the activities originally planned for the first two years of research of the Domain 1 and Domain 3 Task Forces have been reassigned to a new research unit, the Modeling Cross-domain Research Team....[7]

Of the planned outcomes—"prototypes of appraisal and preservation systems, activity models, and guidelines for records preservers"[8]—only the concept of the guidelines remained.

One final problem affected the researchers in Domain 3. Despite the large number of participants in the Project, archivists were in short supply. There were nine Working Groups,[9] joining either as three focus research units or as three domain research units and four cross-domain research units, requiring archival representation with experience from across the spectrum—from small to large archives, from corporate and governmental organizations, from independent institutions and those attached to large parent organizations, as well as archivists with experience with both private-sector and public-sector records, and the artistic, scientific and governmental sectors. Archivists were also required on each of the twenty-three approved case studies and eleven general studies. There were, in fact, not enough archivists to go around, and this problem was exacerbated by the difficulty that a number of the participating archival institutions seemed to have in maintaining consistent representation to the Project over the long term.

In the spring of 2005, with a number of case studies and general studies completed and with the Chain of Preservation model lacking only the narrative report, members of Domain 3 finally acknowledged that, in InterPARES 2 as in life, appraisal and preservation come last. It was now time for Domain 3 to begin its work.

The InterPARES 2 research proposal described the appraisal function, emphasizing the variations introduced when an appraisal addresses records in digital, rather than analogue, form:

> Appraisal assesses the continuing value of the records but it also assembles evidence for the presumption of their authenticity, and identifies the digital

---

[6] See http://www.interpares.org/ip2/ip2_midterm_progress.cfm.
[7] Luciana Duranti (2004), "InterPARES 2 Project Midterm Report to the Social Sciences and Humanities Research Council." MCRI Grant No. 412-2001-1003, 4 (unpublished).
[8] InterPARES 2 Project, Domain 3 Web page. Available at http://www.interpares.org/ip2/ip2_domain3.cfm.
[9] For a summary of the intellectual organization of the Project, see http://www.interpares.org/ip2/ip2_intellectual_organization.cfm.

> components or objects that need to be stored and reproduced to ensure the preservation of authentic records. Appraisal also establishes the feasibility of preserving a given body of electronic records in light of the existing and expected preservation capabilities of the preserver.[10]

This definition reflects the findings of the first InterPARES Project, which highlighted the need for three important shifts in traditional appraisal methodology when applied in a digital environment.

First, there is a need to overtly document the evidence in support of the authenticity of the records, given the ease with which digital records can be accidentally damaged or intentionally modified. With analogue records, the presumption of authenticity is frequently assumed, particularly when they are acquired directly from the creator, which confirms provenance and forms an unbroken chain of custody from creator to preserver.

Second, the definition cited above emphasizes the concept of "digital components," the various bits and pieces that must be identified and preserved to reproduce the complete record. In a vast majority of cases with analogue records, the content of the record and its structure cannot be altered or separated from the carrier—the components of the record are inextricably linked. Much of the context of the record's creation may also be overtly incorporated into this unalterable whole, primarily through the use of file classification numbers on the record and the presence or absence of elements of documentary form.

The third shift addresses the complex preservation alternatives that must now be assessed and costed before a preserver can realistically commit to the long-term preservation of digital records. The range of digital preservation strategies has moved far beyond the adoption of acid-free folders and boxes, although one must acknowledge that even with analogue records, "progress" has offered a steady supply of increasingly difficult formats requiring preservation—from acidic paper, to nitrate still and moving image negatives, to thermal fax paper, to magnetic audiotape.

As the InterPARES 2 research unit responsible for investigating the methods of appraisal and preservation, Domain 3 was tasked with investigating whether the concepts developed by both the Appraisal Task Force and the Preservation Task Force of InterPARES 1 would still apply in the digital environments examined by InterPARES 2. These dynamic, interactive and/or experiential environments could potentially produce records that have no obvious equivalent in the traditional analogue world.

The final report of the InterPARES 1 Appraisal Task Force offered a number of suggestions regarding the practice of appraisal in a digital environment.[11] Based on the analysis of case studies that focused primarily on databases and on document and records management systems, and based on the development of a model of the appraisal function, the Appraisal Task Force made the following conclusions:

- The appraisal of digital records is best conducted when the records are still active. The appraisal of digital records early in their lifecycle greatly improves the documentation available to the archivist about the operational role of the records in the creator's organization and provides technical information about how the application generates and

[10] Luciana Duranti (2001), "International Research on Permanent Authentic Records in Electronic Systems (InterPARES): Experiential, Interactive and Dynamic Records," SSHRC MCRI InterPARES 2 Project Proposal, 412-2001, 1.1-12. Available at http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf.

[11] See Heather MacNeil et al., "Part One – Establishing and Maintaining Trust in Electronic Records: Authenticity Task Force Report," in Duranti, *Long-term Preservation*, op. cit., 19–65. Online reprint available at http://www.interpares.org/book/interpares_book_d_part1.pdf.

maintains records during the active and semiactive periods. Early identification of records with archival value should improve the chances that these records will not be destroyed accidentally or fall into unrecoverable technological obsolescence.

- The medium of records affects the process of appraisal but not the fundamental task of assigning value. Tasks such as the formal identification of indicators of authenticity,[12] the assessment of preservation strategies and the ongoing monitoring of appraisal decisions all represent new or expanded tasks in the appraisal process.
- Monitoring the appraisal decision to confirm the continued archival value of selected records is a necessary activity in the digital environment. The concept that appraisal decisions need to be re-visited at regular intervals is particularly applicable when appraisals are conducted, as suggested above, some time before the actual transfer of inactive records will occur, and to keep pace with the rapid rate of technological change.
- Information compiled during appraisal must be "packaged" and carried forward to assist with ongoing monitoring, transfer, processing, description, preservation and subsequent access. The automation of all aspects of archival work, including appraisal, will greatly facilitate this ongoing re-use, at subsequent phases, of information collected during the appraisal process.

Following a similar methodology, the InterPARES 1 Preservation Task Force, in its final report,[13] made the following conclusions:

- It is not possible to preserve a digital record: it is only possible to preserve the ability to reproduce the record. As with the findings related to appraisal, this statement emphasizes the concept of "components" when discussing digital records. This concept is not unknown among analogue technologies—examples of simpler forms of the concept include the negative and the print in photography, or the negative and positive image, the optical and/or magnetic soundtrack, the composite print and outtakes in moving images. Digital records offer the most complex version of the component system, requiring careful attention to multiple dependencies related to hardware, operating systems and application software.
- The intellectual and physical components of a digital record do not necessarily coincide; a digital component is distinct from an element of documentary form. For example, the content of a record may include both text contained in a word processing file and a table generated by spreadsheet software. Technically, the text file may only contain a link to the spreadsheet file, which in turn may depend on the spreadsheet software rather than word processing software to display it by recognizing and actualizing formatting information.
- The process of preservation must be thoroughly documented as a primary means for protecting and assessing authenticity over the long term. Since the process of preservation begins at creation, responsibility for this thorough documentation rests with both the creator and the preserver. In the past, the stability of most analogue record forms frequently allowed creators to ignore preservation concerns until the inactive records were transferred to the preserver.

[12] See Authenticity Task Force (2002), "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records," ibid., 204–219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf.
[13] See Kenneth Thibodeau et al., "Part Three – Trusting to Time: Preserving Authentic Records in the Long Term: Preservation Task Force Report," ibid., 99–116. Online reprint available at http://www.interpares.org/book/interpares_book_f_part3.pdf.

When InterPARES 2 researchers developed the various research methodologies that would be used during the course of the Project, the overall thrust was to discover whether the case studies, modeling exercises or surveys planned for InterPARES 2 would uncover any theories or practices that disagreed with the existing recommendations of InterPARES 1. Overall, the findings of InterPARES 1, based on large databases and records management applications, had fit smoothly into existing archival knowledge and practice. Essentially, these databases and applications were found to produce digital manifestations of record types that were well-established in the paper-based recordkeeping environment, such as case files and textual records. Would the study of newer forms of digital records by InterPARES 2 researchers suggest the need for new appraisal criteria, or the addition of steps to current appraisal practices, or the development of new preservation strategies or practices?

Research team

The following is a list of researchers and research assistants who participated in the Domain 3 Task Force throughout the Project.[14]

Chairs and Co-chairs:

Yvette Hackett and Sally Hubbard	Jun 2005 - Dec 2006 (Co-chairs)
Hans Hofman and Sally Hubbard	Feb 2004 - Jun 2005 (Co-chairs)
Ken Thibodeau	Jan 2002 - Dec 2003 (Chair)

Researchers:

Howard Besser	New York University, USA—Working Group 3.1
Ann Butler	New York University, USA—Working Group 3.1
Kevin Glick	Yale University, USA—Working Group 3.2
Elaine Goh	National Archives of Singapore—Working Group 3.3
Yvette Hackett	Library and Archives Canada—Working Group 3.1
Babak Hamidzadeh	Library of Congress, USA—Working Group 3.2
P.C. Hariharan	Systems Engineering & Security, Inc., USA—Working Group 3.2
Ken Hawkins	National Archives and Records Administration, USA—Working Group 3.3
Hans Hofman	National Archives of the Netherlands—Working Group 3.3
Sally Hubbard	Getty Institute, USA—Working Group 3.1
Mary Ide	WGBH, USA—Working Group 3.1
Randal Luckow	Turner Broadcasting, USA—Working Group 3.1
Richard Marciano	San Diego Supercomputer Center, USA—Working Group 3.3
Evelyn McLellan	Insurance Corporation of British Columbia, Canada—Working Group 3.3
Reagan Moore	San Diego Supercomputer Center, USA—Working Group 3.2
Isabella Orefice	Associazione Nazionale Archivistica Italiana—Working Group 3.1
Jim Suderman	Archives of Ontario, Canada—Working Group 3.3
Ken Thibodeau	National Archives and Records Administration, USA—Working Group 3.3

[14] Researcher membership in Domain 3 changed substantially over the five years of the Project. Among those who were interested in appraisal and preservation issues but were unable to participate for the full length of the Project are: Filip Boudrez, City Archives of Antwerp/the DAVID Project, Belgium; Michèle Cloonan, Simmons College, USA; Margaret Hutchison, British Columbia Archives, Canada; Glenn Isaac, British Columbia Archives, Canada; Rick Kopak, The University of British Columbia, Canada; Rich Lysakowski, CENSA, USA; Jean-Stéphen Piché, Library and Archives Canada; Shelby Sanett, Amigos Library Services, Inc., USA; Lynne Tibbitt, British Columbia Archives, Canada; and Bill Underwood, Georgia Tech Research Institute, USA.

James Turner	Université de Montréal, Canada—Working Group 3.1

Research Assistants:

Tom Anderson	The University of British Columbia, Canada
Patsy Baudoin	Simmons College, USA
Carolyn Casenas	The University of British Columbia, Canada
Natalie Catto	The University of British Columbia, Canada
Alan Doyle	The University of British Columbia, Canada
Adam Farrell	The University of British Columbia, Canada
Fiorella Foscarini	The University of British Columbia, Canada
Peggy Heger	The University of British Columbia, Canada
Tracey Krause	The University of British Columbia, Canada
Karen Langley	The University of British Columbia, Canada
Catherine Miller	The University of British Columbia, Canada
Luke Meagher	The University of British Columbia, Canada
Jennifer Meehan	The University of British Columbia, Canada
Shaunna Moore	The University of British Columbia, Canada
Elisheba Muturi	The University of British Columbia, Canada
Carolyn Petrie	The University of British Columbia, Canada
Brian Trembath	The University of British Columbia, Canada
Sherry Xie	The University of British Columbia, Canada

Research questions

The four research questions that were to be answered over the course of the Project were identified in the original Project proposal, as follows:

1. How do the appraisal concepts, methods and models developed by InterPARES 1 for the administrative and legal records created in databases and document management systems apply to the appraisal of the records of artistic, scientific and governmental activities resulting from the use of the technologies examined by InterPARES 2?
2. How do the preservation concepts, methods and models developed by InterPARES 1 for the administrative and legal records created in databases and document management systems apply to the preservation of the records of artistic, scientific and governmental activities resulting from the use of the technologies examined by InterPARES 2?
3. What preservation paradigms can be applied across activities and technologies? What preservation paradigms are required for specific types of records resulting from each activity?
4. What metadata are necessary to support appraisal and preservation of authentic digital records resulting from each activity?

As noted earlier, the model-based analysis originally envisioned for questions 1 and 2 were subsequently transferred to the Modeling Cross-domain. However, the Domain 3 researchers would still be able to recognize any shifts in how appraisal and/or preservation activities might need to be conducted in the artistic, scientific and governmental environments being examined in the InterPARES 2 case studies, as opposed to the more traditional bureaucratic environment targeted in InterPARES 1. Of particular concern here was determining whether appraisal or preservation procedures would require any additional adjustments or deviations to accommodate digital technologies beyond those already identified by InterPARES 1.

The paradigmatic shift encountered in the InterPARES 1 research resulted in a shift in the language of digital preservation, from the concept of the preservation of "the record" to an acknowledgement of the importance of identifying and preserving "digital components" and their relationships.[15] Of particular concern here was determining whether the use of interactive, experiential or dynamic systems would require a similar shift in archivists' understanding of how to carry out archival activities.

Finally, it was determined that the answer to the fourth, metadata question should flow naturally from archivists' review of the case studies. An inability to appraise the case studies' records or to determine an appropriate preservation strategy for them would strongly suggest missing metadata, and should actually identify the type of metadata required.

Research Methodology

The concept of appraisal, as practised by archivists, is largely unknown among records creators. The identification of the value of their records as transitory, short term or long term is based on business rules and practices and, in some professions, legal requirements. But these values are generally assigned to records without reference to issues of reliability, authenticity or the value of the records from an historical, cultural or social perspective or the future interests of any potential third-party researchers. Therefore, the initial meetings of the Domain 3 researchers concentrated on the development of methodologies appropriate to the study of the creators' maintenance activities, despite the differences between record maintenance strategies and the time frames or concerns involved in archival preservation.

The first research initiative to be considered by the Domain 3 researchers was the development of a bibliography on digital preservation. This was eventually rejected, since there are many excellent bibliographies already available, and they are well documented by the National Library of Australia's Preserving Access to Digital Information (PADI) site.[16]

Case studies

> *The Domain 3 task force will use the knowledge gained in the course of the case studies...*[17]

The Domain 3 researchers considered the usefulness of initiating one or more case studies designed to answer the research questions. This would have involved studying organizations with appraisal and preservation either as their core activity, or with internal units tasked with appraisal and preservation on behalf of the larger organization. The archivists in Domain 3 were well aware of the fact that, internationally, very few archival units had developed practices and procedures in these areas and, of those that had, most had already been targeted by other research units within the Project, be they a focus, domain or cross-domain.

The Domain 3 researchers confirmed the value of participating in already approved InterPARES 2 case studies and surveys and of analyzing those results from an appraisal and preservation perspective, rather than launching competing projects.

---

[15] See Thibodeau et al., "Preservation Task Force Report," op. cit., 6–7.

[16] The subsection of the PADI site devoted to bibliographies contains resources about preserving access to digital information (see http://www.nla.gov.au/padi/format/bib.html). Recently, the indicator "Historical" was added to identify less current material.

[17] InterPARES 2 Project, Domain 3 Web page. Available at http://www.interpares.org/ip2/ip2_domain3.cfm.

Summaries of the case studies are available on the InterPARES 2 Web site[18] and in the Focus Task Force Report.

Case study interview questions

An early activity of the Domain 3 researchers was to design the following question for inclusion in the suggested list of case study questions:[19]

> *39. Do you have a standard procedure when it is time to preserve [your documents/work] for the long-term?*

It was hoped that this question would elicit information about any practices implemented or lessons learned about preservation by creators handling even fairly recent digital objects. The limited perspective of many of those who work in the information technology industry guarantees that obsolescence and interoperability problems can be experienced within extremely short time periods.

A number of other questions were also reviewed, since it was possible that case study respondents would provide information about various aspects of their preservation activities in any number of the suggested questions. Question 34 looked for any efforts at standardization in the work-flow that could represent the traditional concept of the "normal course of business." A broad question, it also allowed the identification of any standards adopted by the records creator:

> *34. Have you had to make rules, or adopt standards to help you in your work? Do you find you have to update them regularly?*

Although primarily a question for the Policy Cross-domain, question 46 addressed the fact that the identification of legal or ethical issues related to the work could have an important influence on the preservation strategies appropriate for the digital objects of a particular case study:

> *46. Do any legal or ethical issues arise from your electronic work?*

Questions 9 and 10 directly addressed the concerns of the Description Cross-domain. The existence of identity and integrity metadata describing a digital object is a *benchmark requirement* for authenticity.[20] Therefore, the descriptive practices of the records creators involved in the case study were an important aspect of their successful preservation:

> *9. Did you create or adopt a standard list of information which you try to record about each file, or work?*
>
> *10. Where did you get it? Do you know if others use the same one?*

Since the process of preservation of digital records really begins at the moment of creation, any number of the suggested questions could also elicit information of interest to Domain 3's research concerns. As case studies were completed, the results of the interviews were reviewed, both from the perspective of the case study questions and the Domain 3 research questions.

[18] See http://www.interpares.org/ip2/ip2_case_studies.cfm.
[19] Domain 3 (2003), "List of possible case study questions that the researchers may ask the subjects of their case studies to acquire the information necessary to answer the 23 questions." Available at http://www.interpares.org/display_file.cfm?doc=ip2_possible_questions_for_interviewees.pdf.
[20] Specifically, Benchmark Requirement A.1 - Expression of Record Attributes and Linkage to Record. See Authenticity Task Force Report, "Appendix 2," op. cit., 210–211. See Appendix 21a for an abridged version of the benchmark requirements.

Domain 3 Template for Case Studies Analysis

During the February 2005 InterPARES workshop in Vancouver, the Domain 3 researchers proposed developing a template for analyzing the Domain 3-related data being generated by the case studies. The template included nine questions related to the appraisal and preservation issues being investigated by Domain 3.[21] The questions were designed to extract information contextualized to fit the template from each of the case studies. Although the information the Domain 3 researchers were looking for about appraisal and preservation activities was largely embedded within sections E (narrative answers to the core research questions) and F (narrative answers to the applicable domain and cross-domain research questions) of the final case study reports,[22] supplementary information could come from interim case study reports or from other available documentation, such as interview transcripts or the creators' Web sites. In some cases, additional information could be inferred, based on researchers' knowledge of certain fields. In addition to the nine questions, the template included instructions and guidelines for answering the questions, along with two appendices that provided detailed descriptions and definitions of key concepts, as well as immediate access to the relevant InterPARES 1 benchmark and baseline requirements and commentaries.

After being presented and discussed by InterPARES 2 researchers at the Chicago workshop in September 2005, a test "walkthrough" of the template was conducted with several case studies at both the Chicago workshop and the International Team meeting in Venice in December 2005. Following revisions and final approval of the template at the February 2006 workshop in Vancouver, the template was applied to the remaining case studies by a team of University of British Columbia research assistants under the direction of the Domain 3 researchers. Once completed, the templates for each case study were forwarded to the appropriate case study principal investigator(s) for review, comment and, as necessary, revision to ensure that the templates had drawn out the relevant information from the case studies. Once validated, the completed templates were analyzed by researchers who represented a wide range of disciplines.

At the end of this process, the Domain 3 Task Force had a tool that assisted the research in several ways. First, it augmented the validation of the case studies that was being undertaken by the Domain 1 Task Force, by highlighting certain gaps in the existing reports and serving as an informal methodological guide for writing the narrative answers to the Domain 3 questions in the final reports of the case studies that were still underway. Second, this tool served as an index to appraisal- and preservation-related information in the case studies, which InterPARES researchers could use to: (a) associate with the analyses done by other research units, including the diplomatic analyses, case study characterizations and modeling walkthroughs; and (b) return to the source material as needed to clear up a given question. Third, by characterizing all of the case studies in relation to the Domain 3 questions, it was possible to identify activities that would affect the ability of archivists to subsequently appraise and preserve the records of the creators in each of the case studies. Finally, this tool provided a more formalized and analytically rigorous compendium of case study-related information upon which to draw while developing the *Preserver Guidelines*.[23]

[21] Appendix 13.
[22] See Appendix 9.
[23] See Appendix 21. The Guidelines also are available in booklet form at http://www.interpares.org/ip2/display_file.cfm?doc=ip2(pub)preserver_guidelines_booklet.pdf.

General studies

In addition to the information extracted from the case studies, the Domain 3 researchers determined that a number of the Project's general studies provided supplementary information that could be used to help address the Domain's research questions and further inform development of the Guidelines for Preservers. These studies and their relevance to Domain 3 are summarized as follows.

General study 01: Persistent Archives Based on Data Grids

The first of the preservation-oriented general studies to receive approval in InterPARES 2, entitled *Persistent Archives Based on Data Grids*, was conducted by Reagan W. Moore of the San Diego Supercomputer Center. Beginning from the premise that preservation environments for digital records are successful when they can separate the digital record from any dependence on the original creating infrastructure, the study investigated the potential use of data grid technology, which supports the management of distributed records. The final report, *Building Preservation Environments with Data Grid Technology*,[24] examined the minimum capabilities required to preserve records, focusing on selected digital holdings of the U.S. National Archives and Records Administration and a number of other data grid implementations around the world.

General study 04: Survey of Recordkeeping Practices of Composers

This is one of three general studies that was derived from the case studies and that involved surveying a group of records creators who were also the subject of a case study.[25] As the Focus Task Force Report explains:

> ...while case study 13 was delving deeply into the work of one composer, Keith Hamel, and the technical details of one specific composition, *Obsessed Again...*, the Focus 1 researchers understood that this single case was not necessarily representative of the full range of adoption and use of digital technologies among composers...[26]

This 2003 survey confirmed that most composers using digital technology work with off-the-shelf commercial software, and almost half of the respondents (forty-seven percent, or seventy-five individuals) have already lost files they considered valuable, through either hardware or software obsolescence.[27] The majority of composers keep digital files primarily for re-use rather than from any concern about long-term preservation or posterity, and thus their primary concern is accuracy, which is clearly related to their need for continued access to the files rather than to concerns about authenticity. Composers work alone; therefore, there is limited access to their digital files by others. This finding also identified the appraisal and preservation issues surrounding the difference between the score and the performance, which are discussed extensively in the report of the task force on Records Reliability, Accuracy and Authenticity (i.e., Domain 2). The only "preservation" measures that the surveyed composers consistently took to

[24] Reagan W. Moore (2004), "InterPARES 2 Project - General Study 01 Final Report: Building Preservation Environments with Data Grid Technology." Available at http://www.interpares.org/display_file.cfm?doc=gs01_final_report.pdf.
[25] General studies 07 (Survey of Recordkeeping Practices of Photographers using Digital Technology) and 09 (Digital Recordkeeping Practices of GIS Archaeologists), discussed in the following two sections, were the other two studies. All three of these general study surveys—dealing with composers, photographers and archaeologists—contained questions relating to preservation practices.
[26] Focus Task Force Report, 7.
[27] Michael Longton (2004), "InterPARES 2 Project - General Study 04 Final Report: Survey of Recordkeeping Practices of Composers." Available at http://www.interpares.org/display_file.cfm?doc=gs04_final_report.pdf.

protect their files were either to produce a backup copy on an external medium or else to upgrade file formats to more current versions.

General study 07: Survey of Recordkeeping Practices of Photographers Working with Digital Materials

This 2004 survey was completed by a larger number of practising photographers (*n*=402) than that of the composers (*n*=161) and revealed a rather different situation:

> In response to the pressures of faster turnaround times, creative innovation and remote transmission, professional photographers have universally embraced the transition from analogue to digital photography. The majority of photographers identified their practice as completely digital, allocating the use of analogue film to the occasional personal project. Even amongst those who identified their practice as a hybrid of digital and analogue, the bulk of their images were born digital - only a small percentage of analogue images were made and most of these were eventually digitized.[28]

Like composers, photographers were most interested in maintaining their digital images to support re-use and reference, although the business aspects of their profession, such as journalism or forensic work, had made them more sensitive to issues such as authenticity and copyright. Their work environment also made many of them aware of transmission problems with digital files, as well as hardware, software and storage issues. In addition to refreshing digital media and upgrading file formats, the significant number of photographers who responded to the survey routinely record metadata, practice quality control, implement security procedures and maintain multiple versions of their digital images, including in-camera file formats and working drafts and versions.

General study 09: Survey of the Digital Recordkeeping Practices of GIS Archaeologists Worldwide

This 2004 survey of GIS use by archaeologists examined one field of scientific endeavour. Unlike the previous two surveys, this survey could compare its results against a 1998 survey investigating similar issues.[29] The InterPARES survey revealed that a significant increase in the awareness of archaeologists about digital issues had occurred during the intervening six years since the 1998 survey:

> A key indicator of this increased awareness is the growing sense of frustration expressed by many participants over the current lack of suitable long-term preservation repositories available to archaeologists, as well as over the continuing absence of any concerted, profession-wide response to these particular issues and concerns.[30]

On the other hand, the survey also revealed that many respondents engaged in idiosyncratic and ad hoc file creation, management, preservation and/or documentation practices. Lack of training in GIS software was cited as one problem, as was a lack of knowledge about how to

---

[28] Jessica Bushey and Marta Braun (2006), "InterPARES 2 Project - General Study 07 Final Report: Survey of Recordkeeping Practices of Photographers using Digital Technology," 30. Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs07_final_report.pdf.

[29] Khalid Gourad. "Geographic Information Systems in Archaeology: A Survey" (unpublished Master's thesis, Hunter College of the City University of New York, Department of Anthropology, 1999), 75 pp. Both the thesis and the survey are available at: http://khalid.gourad.com/thesis/.

[30] Randy Preston (2006), "InterPARES 2 Project - General Study 09 Final Report: Digital Recordkeeping Practices of GIS Archaeologists Worldwide: Results of a Web-based Survey," 3. Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs09_final_report.pdf.

provide adequate documentation within the bounds of available project funding. The specialized nature and complexity of GIS software was reminiscent of the composers' situation, while the problem of ownership of the digital files (the archaeologist or the client/funding source) resembled the photographers' problems with employers, freelancing and commissions.

General study 10: Preservation Practices of Scientific Data Portals

This general study involved a researcher survey of thirty-two data portals from a variety of scientific disciplines designed to reflect the heterogeneity of scientific research. Structured information was collected about each portal's service, its host institution, income sources, access fees and metadata standards. In addition, information about preservation practices, as well as statements related to the data quality, accuracy and/or reliability of each portal's data were collected. The purpose of the study was to compile structured information about the standards and protocols in place at science data portals (a.k.a., archives, repositories, catalogues, etc.) that would, in turn, provide better insight into how the key InterPARES 2 issues of accuracy, reliability and authenticity are interpreted and understood in the sciences and how this understanding is used to underwrite confidence in data accuracy, reliability and authenticity as practiced and implemented in the natural and physical sciences.

General study 11: Selecting Digital File Formats for Long-term Preservation

This general study, which focused on the criteria used by various preserver organizations to select digital file formats for preservation, was originally developed in the Policy Cross-domain. The initial interest was in studying policy statements related to preservation activities, but the research was transferred to Domain 3 when it began to focus on the more specialized question of criteria for format selection. The final report's seven recommendations address issues ranging from vague or misleading terminology, to the need to clearly distinguish between various file types, to compression, among others.[31]

General study 05: An Examination of the Processes to Preserve and Manage Electronic Records

This general study, which was never completed, would have constituted the third part of an extended research project initiated during InterPARES 1. Begun in 2000, the initial Survey of Preservation Practices and Plans investigated the current state of digital preservation practices and plans at thirteen institutions, projects or programs in North America, Australia and Europe.[32] The findings, reported in June 2001, revealed:

> ...a number of preservation techniques were in use but that none of them could be considered meeting archival requirements for authenticity. The study also revealed that while developing technological processes to preserve authentic electronic records, almost every institution had deferred costing digital preservation processes and implementing digital preservation policies.[33]

---

[31] Evelyn Peters McLellan (2006), "InterPARES 2 Project - General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation," 16–17. Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_english.pdf. French language version available at http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_french.pdf.

[32] Michèle V. Cloonan and Shelby Sanett (2002), "Where We Are Now: Obliquity and Squint?" *The American Archivist* 65(1): 70–106. The June 2001 draft of the report, entitled "Survey of Preservation Practices and Plans," is available as a draft appendix to the Preservation Task Force's final report at http://www.interpares.org/display_file.cfm?doc=ip1_survey_of_preservation_practices-plans.pdf.

[33] Michèle V. Cloonan and Shelby Sanett (2004), "InterPARES 2 Project - General Study 05 Final Report, Round 2: The Preservation of Authentic Electronic Records: *Ad hoc*, Inconsistent, or Strategic?" 3. Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs05_r2_final_report.pdf.

The second phase of the survey was undertaken between August 2001 and February 2003, thus bridging the transition from InterPARES 1 to InterPARES 2. The authors re-surveyed eight of the thirteen organizations that participated in the first round, and conducted eighteen in-depth interviews with selected key informants. This process provided a more detailed understanding of the preservation strategies being explored and measured the progress that had been made in the interim. They found that the programs' successes were uneven, moving ahead in some areas while lagging behind in others:

> For example, institutions are beginning to think about cost issues and models, but have been slow to develop digital preservation policies and plans. As one of our survey respondents in round 2 observed about his own institution, "As long as there is no plan, the risk will be that preservation will be ad hoc, inconsistent, and not imbedded in the organization."[34]

The third survey—An Examination of the Processes to Preserve and Manage Electronic Records: Round Three at The National Archives of Australia and WGBH—was approved by InterPARES 2 in February 2004. The intent was to conduct in-depth studies of two organizations that had been surveyed in the two previous rounds. One particular focus would be policy development and planning, which had been identified as lacking in the previous report. The second area of emphasis would be the specific preservation strategies adopted by the two organizations. Unfortunately, issues of distance and availability—of both the researchers and the subjects of the research—eventually posed insurmountable problems and the final phase of the project could not be carried out.

Research Findings

Analysis of the case and general study data suggests that too many records creators are still neglecting the long-term preservation of their digital files, whether they be static or dynamic, evidential or experiential, historically significant or interactive. The range and breadth of the many case and general studies does mean that almost every variation in approach and attitude was documented at least once. However, some distinctions can be made between individuals and small organizations on the one hand and large corporations and government institutions on the other. The InterPARES 2 case studies done by the government focus provided similar findings to those done in InterPARES 1, demonstrating a greater knowledge and awareness of appropriate practices for records creation and maintenance, whether of digital or paper-based records. Unlike individuals and small organizations, the larger institutions also tended to have analogue recordkeeping systems in place, which the institutions could fall back on by identifying the printed paper copy as the authoritative record and moving it into the existing analogue records management system. One exception to this observation occurred with Web sites, which both large and small organizations frequently treated almost as ephemera—insofar as organizations made little or no effort to set aside and preserve fixed versions of their Web sites—although this was balanced by those organizations that viewed Web sites as either a type of "recordkeeping system"[35] or as a type of legacy site on which to preserve their work.[36]

[34] Ibid., 3–4.
[35] See, for example: Henry Daniel and Cara Payne (2004), "InterPARES 2 Project - Case Study 02 Final Report: Performance Artist Stelarc," 7–13. Available at http://www.interpares.org/display_file.cfm?doc=cs02_final_report.pdf.
[36] See, for example: Martine Cardin (2004), "InterPARES 2 Project - Case Study 01 Final Report: Arbo Cyber, théâtre (?)." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_report_english.pdf. Original French language version available at http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_report_french.pdf.

With individuals and small organizations, in far too many instances, there were no preservation practices to study and little evidence that even basic maintenance strategies were being pursued with any consistency.[37] In many cases, the maintenance procedures that were in place were more the result of a happy accident than of an intentional act of preservation. The most popular strategy was undoubtedly "redundancy," or making a copy somewhere other than on the creator's computer or server hard drive. For large organizations, regular backups run by information technology staff were and are now standard practice.[38] For smaller organizations and individual users, some form of backup was frequently cited.[39] One improvement in the redundancy strategy is the disappearance of the old DOS-based concept of "backup" formats, which were highly sensitive to upgrades in the operating system. Steady decreases in the cost of storage means more copies are basic copies; that is, copies saved in the file formats in which the records were originally created, or in which they were last used and saved, thus (theoretically) making them more immediately human-readable in the creator's usual desktop environment. Directory structures and file names are immediately accessible without the need to apply an extra "Undo" command to re-open the—often compressed—backup or protection copy. Where very large files are concerned, use of compression may reintroduce the extra layer of software needed to re-access the backup copy of the records.

These redundant copies were also, in some cases, stored in off-site locations. Perhaps the most comprehensive example of this approach is found in case study 21 (Electronic Filing System (EFS) from the Supreme Court in Singapore), where it is noted:

> There is both a daily and weekly backup of data, which are kept in an off-site location to ensure the full restoration of data in the event that the system fails. For example, every time a record is filed by the law firm, the records are stored on three disks. One disk is stored permanently in the jukebox to facilitate online access to information. The second disk is taken to an off-site storage at the end of the week and the third disk is sent off-site once the disk is full.[40]

In some cases, redundant copies were distributed between the creator's home and office,[41] neither of which offered stable environmental conditions such as temperature, relative humidity

[37] Eight basic maintenance strategies are described in an appendix to the *Preserver Guidelines* for developed by Domain 3 researchers and appended to this report (see Appendix 21c). In brief, these strategies are: (1) clear allocation of responsibilities; (2) provision of the appropriate technical infrastructure; (3) system maintenance, support and replacement; (4) transfer of data to new storage media on a regular basis; (5) adherence to appropriate conditions for storage media; (6) redundancy and geographic location; (7) system security; and (8) disaster planning.

[38] See, for example, in the arts focus, James Turner, et al. (2004), "InterPARES 2 Project - Case Study 09(3) Final Report: Digital Moving Images - Commercial Film Studio," 8–10. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_final_report.pdf; in the science focus, Tracey P. Lauriault and Yvette Hackett (2005), "InterPARES 2 Project - Case Study 06 Final Report: Cybercartographic Atlas of Antarctica," 17. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs06_final_report.zip; and, in the government focus, Elaine Goh (2005), "InterPARES 2 Project - Case Study 21 Final Report: The Electronic Filing System (EFS) of the Supreme Court of Singapore," 32. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs21_final_report.pdf.

[39] See, for example, reliance on backup to: (1) CD-ROM or DVD in Daniel and Payne, "Case Study 02 Final Report," op. cit., 13, and Bart Ballaux (2005), "InterPARES 2 Project - Case Study 26 Final Report: MOST Satellite Mission - Preservation of Space Telescope Data," 7. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs26_final_report.pdf; and (2) separate computer(s) in Sydney Fels and Seth Dalby (2004), "InterPARES 2 Project - Case Study 15 Final Report: *Waking Dream*," 4. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs15_final_report.pdf. See also Ballaux, "Case Study 26 Final Report," op. cit.

[40] Goh, "Case Study 21 Final Report," op. cit. See also Ballaux, "Case Study 26 Final Report," op. cit., for an example of the use of off-site backup procedures by a creator in the science focus.

[41] See, for example, the strategy used by the creator in case study 14, in which copies of the GIS data and related records are "maintained," in an ad hoc manner, "on either the creator's personal computer, the organization's server and/or copied onto CD-ROMs" (Richard Pearce-Moses, Erin O'Meara and Randy Preston (2004), "InterPARES 2 Project - Case Study 14 Final Report:

or clean air. Moreover, although this type of "informal" distributed storage strategy does offer increased protection against vandalism, theft or fire, it may be of little use in larger-scale disasters such as earthquakes or floods.[42] Aside from the limited data recovery functionality provided by the relatively basic, and often ad hoc, practices noted above involving *selective backups* of data or files,[43] formal and explicit disaster planning was almost non-existent among all the creators examined.[44]

A number of the case studies and surveys also confirmed what the Domain 3 researchers already knew based largely on their own records creation and maintenance practices. For example, most creators learn only the minimum required about a technology to get it working and perform basic functions. Few users become experts or learn to appreciate the implications for long-term records preservation of, for example, configuration choices, or add-on programs generated by third-parties, or relevant functionality within their chosen software environment.[45]

The Domain 3 researchers observed that, especially in the arts, hardware components, particularly peripherals, can play an essential role in the accurate and authentic reproduction of, for example, a work of art.[46] However, hardware dependencies can be almost impossible to overcome, especially if responsibility for preservation is transferred to a trusted custodian who lacks access to the required hardware and/or the ability or resources to maintain such hardware. Acknowledgement of this fact led the Domain 2 researchers to recommend that, whenever possible, records creators should strive to "eliminate dependence on hardware by transferring hardware functionalities to software (i.e., use a software application to simulate the actions of a

---

Archaeological Records in a Geographical Information System: Research in the American Southwest," 28. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs14_final_report.pdf.

[42] Of course, the same concern may also apply in more "formal" distributed storage arrangements, such as the one noted above in case study 21, depending on the nature of the off-site storage facilities used and their geographic proximity to the on-site location.

[43] This is in contrast to *comprehensive system backups* in which the operating system, all software applications and all data/files are backed up so that the *entire system* can be restored in the event of a large-scale system failure (see *Creator Guidelines*, guideline 8.A, in Appendix 20). Two notable exceptions to the otherwise exclusive reliance on selective backup strategies were observed in the case studies; one in the science focus, the other in the government focus. In case study 26 (MOST Satellite Mission), it is stated that, "[s]ince the MOST researchers mostly work with custom-made software, one of the issues related to their work is the preservation of the software. Backups are made of the various programs that are used in the project. Moreover, if anything is changed (i.e., added) to one of the programs, the old version of the software is always preserved. In this way, the researchers are always capable of recreating the results that they created previously" (Ballaux, "Case Study 26 Final Report," op. cit., 7–8). In case study 12 (Antarctic Treaty Searchable Database), it is noted that "the digital-record entities in the *Antarctic Treaty Searchable Database* have been maintained on servers with backup copies on additional hard-drives as well as on webCDservers ... which replicate the full functionality and contents of the website.... In one instance, the webCDserver was used to restore the website for the *Antarctic Treaty Searchable Database*" (Paul Arthur Berkman et al. (2005), "InterPARES 2 Project - Case Study 12 Final Report: Antarctic Treaty Searchable Database," 39. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs12_final_report.pdf). Although reference to the backed up entities in the foregoing quote is limited to "the digital-record entities," it seems apparent that, in fact, copies of the operating system and the *Digital Integration System* software application are also included on the backup webCDservers, which the creator elsewhere describes as containing "fully executable copies of the websites" (Ibid., 18).

[44] Case study 09(3) (Commercial Film Studio) provides the only explicit reference to disaster planning (see Turner et al., "Case Study 09(3) Final Report," op. cit., 15 (see answer to question 3.2).

[45] An example of this is provided in general study 09 (Digital Recordkeeping Practices of GIS Archaeologists), where it is noted that "several of the [survey] participants admitted feeling somewhat overwhelmed by the analytical potential of their GISs, due, primarily, to a lack of formal GIS training." As the report cautions, "[i]f true, this could potentially have ramifications with regard to the broader issue of the long-term preservation of digital GIS records. In fact, it is possible that the more alienation, disconnection and/or intimidation one feels toward one's GIS projects, the less likely one may be to feel the sense of obligation, initiative and/or competence necessary to effectively address the project's long-term preservation requirements" (Preston, "General Study 09 Final Report," op. cit., 32).

[46] See, for example, the discussion in case study 13 (*Obsessed Again...*) of an attempt by the Domain 2 researchers to resurrect an electroacoustic work that had already fallen victim to technological obsolescence. A more concise summary of this particular activity is also provided in the section titled "A strategy for preventing technological obsolescence of an artistic work" of the Domain 2 Task Force Report. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_3_domain2_task_force.pdf.

piece of hardware," [since] "[t]his provides a more stable way to retain the function when the hardware becomes obsolete."[47]

Another observation noted by the researchers is that a number of creators are simply ignoring preservation issues, although here too the reasons are varied. In some cases, creators (especially those in the arts focus, but also in the science focus) noted that it would be easier and less expensive to re-create some types of content in digital form at a later date than it would be to attempt to carry forward and convert or provide an emulation platform for highly proprietary material. This view is best epitomized in the following excerpt from the arts focus case study 09(3) (Commercial Film Studio):

> Since important hardware and software changes have usually occurred since the artwork was created, in the environment we studied it is deemed more economical to re-create artwork if it is to be reused than to engage in a process of migrating or otherwise upgrading it in case it will be reused. This approach limits the amount of time, energy and money spent on long-term preservation of digital entities and avoids the need to implement preservation strategies that respond to problems of hardware and software obsolescence.[48]

A similar sentiment was expressed by the creator in the science focus case study 14 (Archaeological Records in a Geographical Information System), who, in response to a question about the intermediary files (e.g., cost surface algorithms) he creates and their relationship to his overall GIS research, stated:

> I guess I don't ascribe a lot of value to those intermediary documents for the most part, because they're easily recreated ... Whenever I'm trying to teach somebody about GIS, people will say, here's the data I have and here's what I want to get to. And the first [thing] that I always [say] – my first caveat that I offer them is, if you ask 10 different GIS people, they'll tell you 10 different ways to get there. Each one of those different ways will produce different intermediary files, and so I don't think, for the most part, that those things have much value. Like I said, they're easy to reproduce, if I did want to, I could always make another slope file, if I lose one or something.[49]

In other cases, some artists do not want their works preserved over the long term if doing so might compromise certain characteristics that they consider essential to the essence of their artworks, such as the ephemerality and variability of works that change each time they are reproduced. An example of this is provided in case study 15 (*Waking Dream*), which examines a multimedia performance art piece. One of the performers and co-creators of the work, Sachiyo Takahashi, "views *Waking Dream* as a performance art piece defined by her role as performer," a view that she believes "precludes performances of the piece without her involvement [as one of the performers]."[50] Based on this interpretation, it would be impossible to "preserve" the ability to reproduce the work or, at least, an authentic reproduction of the work beyond the performer's death.

These observations led InterPARES researchers to further explore and refine their understanding of the distinction between digital documents that are either works or that *document* specific performance events, versus digital objects that either *enable* subsequent reproduction of a work or performance—a concept more closely analogous to a "negative" in

[47] See *Creator Guidelines*, guideline 9, in Appendix 20.
[48] Turner et al., "Case Study 09(3) Final Report," op. cit., 4.
[49] Pearce-Moses et al., "Case Study 14 Final Report," op. cit., 31.
[50] Fels and Dalby, "Case Study 15 Final Report," op. cit., 8.

photography—or provide *instructions* about executing a work or performance—a concept more closely analogous to a "score" in music—with the added fact that these *enabling* and *instructive records* will actually play an active or an instructive role in the re-instantiation of the work or performance. In contrast to instructive records, which "are intended to be read by humans and, therefore, are materialized by being reproduced from stored digital components into a human-readable form,"[51] a key characteristic of enabling records, whether in the arts[52] or in the manufacturing sector where other examples were found,[53] is that they contain instructions intended to be read by a machine and thus

> …achieve their purpose in the digital form in which they are stored[64] and, conversely, cannot achieve that effect if transformed into human-readable format. Moreover, as long as they remain active, enabling records must be maintained in the systems in which they were created—or in systems with identical functionality. Otherwise, they will not produce or enable the interactions, experiences, performances or other processes they were intended to generate.[65] [54]

Still other creators feel they can either delay addressing preservation issues—or ignore them altogether—because they believe that the information technology industry will come up with whatever solution will be needed, whenever they need it. "Trusting the vendor" may work in cases where a technology has been widely adopted and there is an equivalent product on the market to which consumers could move. On the other hand, it is instructive to remember how many physical media since the 8-track tape have been discontinued by manufacturers, even though obsolescence of such media meant, for example, that consumers had to migrate their entire music or video collections to the new technology.

This willingness to "trust the vendor" (or the technology industry in general) seems to be the assumption driving the adoption of digital signature technology despite the well-known problems with carrying this "authenticity" solution forward over time.[55] This, in fact, appears to be the

---

[51] Luciana Duranti and Kenneth Thibodeau (2006), "The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES," *Archival Science* 6(1): 60. (Note: a reprint of this article is included in Appendix 2).

[52] For example, the custom written program code used to control the functionality of a remote control device used by a performer in *Waking Dream* (see Fels and Dalby, "Case Study 15 Final Report," 2 (technological context)), and the various custom written software patches used in *Obsessed Again...*, including a software synthesizer patch of coded instructions that controls the sounds that the computer causes to be played on the synthesizer in response to the notes played by the bassoonist (see J. Scott Amort (2004), "InterPARES 2 Project - Case Study 13 Final Report: *Obsessed Again...*," 3 (see answer to question 4). Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs13_final_report.pdf.

[53] For example, the Pro-Engineer CAD system file used to create the original CAD files in case study 19 (see Kenneth Hawkins (2006), "InterPARES 2 Project - Case Study 19 Final Report: Preservation and Authentication of Electronic Engineering and Manufacturing Records," 6. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs19_final_report.pdf.), meets the definition of enabling record in that it stands as a fixed set of instructions only readable by a machine and that supports subsequent actions and provides instructions to engineering robots, etc. The addition of semantic or actionable metadata to the subsequent iterations of the same fixed information provides the basis for additional subsequent actions; namely, the "interrogation" of the knowledge-enhanced formats by reasoning engines.

[54] Duranti and Thibodeau, "The Concept of Record," op. cit. Note: footnote references in the quote are from the original text, and are not reproduced here. As Duranti and Thibodeau further clarify, "[e]xamples of "enabling" records include software patches that enable a musical instrument to interact with a computer, software in online marketing sites that interprets data about a visitor's actions on the site to determine what elements of content should be presented next to that visitor, and software agents that enable interacting business applications to execute transactions autonomously" (Ibid., 59).

[55] As is emphasized in the *Creator Guidelines*, guideline 6 (see Appendix 20), *technology-dependent* authentication techniques, such as digital signatures, are subject to obsolescence themselves. In fact, by virtue of their purpose and inherent functionality, digital signatures cannot, at present, be migrated to new or updated software applications together with the documents to which they are attached. Moreover, the life of digital signatures and other authentication technologies may be much shorter than the length of time that even a temporary document not requiring migration may need to be maintained, because authentication technology is changing rapidly. A number of different preservation strategies (see Appendix 21c, Part B) would eventually require the conversion of a digitally-signed record to a new logical format. Following this process, it is unlikely that the

situation in case study 18 (Computerization of Alsace-Moselle's Land Registry), where a specifically dedicated administrative body, GILFAM (Groupement pour l'Informatisation du Livre Foncier d'Alsace-Moselle), created in 1994 and charged with overseeing computerization (and digitization) of France's land registry

> …has the legal responsibility to provide continued access to the land registry in a fashion that preserves its evidential value, in conformance with Article 1316-1 of the Civil Code, regardless of technological change, [yet] has not considered the problem of maintaining the digital entities, except through the mechanisms afforded by system vendors when upgrading the database management system. In addition, it has not considered the problem of maintaining the evidential value of digital signatures through technological evolution.[56]

Still other creators feel confident about their ability to "self-monitor" the evolution of technology and standards affecting the long-term preservation of records and to adapt their preservation systems and practices accordingly in time to keep pace with the changes. A good example of this approach is provided in case study 21, where the creator (the Supreme Court of Singapore) describes its current, apparently ad hoc, approach to preservation as consisting of an attempt "to foresee new standards/technology before the current technology that has been deployed becomes fully obsolete," so as to provide "sufficient time to migrate to the new technology/standards without loss of data."[57] Although monitoring the technological environment and taking steps against hardware and software obsolescence are key recommendations provided in the *Creator Guidelines*,[58] the activities must be supported by an integrated framework of policies and procedures to be effective. Unfortunately, this does not appear to be the situation in the case of the Supreme Court of Singapore, since it is acknowledged later in the report that although "[t]he court recognizes that there is a need to address the long-term preservation of electronic records … there is currently no strategy in place *because the court views the system to be fairly current*."[59] The reason provided by the creator in this case for the delay in developing a long-term preservation strategy demonstrates a general concern identified by the Domain 3 researchers in relation to the preservation practices of many of the creators they studied. In fact, beyond a lack of foresight, it shows, more seriously, a fundamental misunderstanding of the temporal scope of the problem at hand and a lack of awareness of the fact that effective long-term preservation begins at creation.

The Domain 3 researchers observed that the technology used by innovators and early adopters, regardless of the focus area in which they belonged, was proprietary[60] and frequently customized. In many cases, the point of the work of these types of creators is to explore, test and push the limits of the available technology, be it hardware or software. This environment cannot

---

document and its signature could be re-validated by the external digital signature infrastructure since the check-sum will have changed. Consequently, creators and preservers are encouraged to rely on *technology-independent* authentication techniques, grounded in effective administrative policies and procedures for records creation, maintenance and preservation, whenever possible.

[56] Jean-François Blanchette, François Banat-Berger and Geneviève Shepherd (2004), "InterPARES 2 Project - Case Study 18 Final Report: Computerization of Alsace-Moselle's Land Registry," 27. Available at http://www.interpares.org/rws/display_file.cfm?doc=ip2_cs18_final_report.pdf.

[57] Goh, "Case Study 21 Final Report," op. cit., 33.

[58] See Appendix 20, guideline 9.

[59] Goh, "Case Study 21 Final Report," op. cit., 38 (emphasis added).

[60] Proprietary either in the sense of commercially available software, the source code of which is privately owned and controlled, or in the sense of custom-built software, the source code of which is developed "in-house" and is not made publicly available. The custom-built software used to process satellite data in case study 26 (MOST Satellite Mission) is a good example of the latter.

wait for internationally agreed-on standards or "open" formats.[61] It requires instead a robust interoperability,[62] at least until such time as widespread adoption of the technology generates some standards, should this ever occur. The Policy Cross-domain, in its final report, noted that "steps taken to ensure interoperability across systems performed many of the same purposes as preservation."[63]

Many scientists view the publication of an article in a journal as a means of preservation, instead of attempting to preserve any of the digital antecedents that led to the published item. This belief was expressed, for example, by many of the survey participants in general study 09 (Digital Recordkeeping Practices of GIS Archaeologists) in response to a question asking them to identify what they considered to be the most important elements or outputs of their GIS projects to preserve for future use or reference, and why. As the study's final report notes, "many participants assume that publication alone constitutes sufficient long-term preservation of their research."[64] In light of this finding, the report goes on to caution that

> …by focusing on the preservation of final reports (and related types of documentation, such as published articles), at the expense of more comprehensive on-going project documentation, archaeologists run the risk of not preserving the types of records required to ensure that the key preservation goals identified by survey participants—i.e., project/data reuse, verification, replication and accountability—can in fact be met.[65]

Government employees may also feel this way about reports that document a study, which they favour for preservation to the underlying data and research that informed the study. This is clearly inadequate, since loss of the data in a manipulable form prevents verification of results, potentially impedes successful duplication of the data-gathering experiment if it is replicable or destroys a particular point in a time series dataset. Moreover, this "publication" argument has lost much of its relevance with the move to electronic journals and reports, where it is clear that publication, even in digital form, cannot replace the original data. The new digital platform for e-journal preservation has created an environment where the related data may well be preserved and can be linked to the publication.

Several case study interviews show that, for many creators of digital objects, there is already a digital equivalent to the concept of trust in the long-term survival of traditional publications: "putting it on the Web." This is, however, a misguided view since there is no organization even attempting to preserve the entire World Wide Web and the millions of changes made to its content every second.[66] In fact, these two proposed preservation strategies are manifestations of

---

[61] For a discussion of the use and understanding of the concept of "file format" in the context of archival institutions, including the terms "standard file format" and "open file format" (also known as "open-source" and "non-proprietary" file format), see McLellan, "General Study 11 Final Report," op. cit.
[62] The InterPARES 2 Terminology Database defines "interoperability" as "the ability of one application/system to communicate or work with another" (http://www.interpares.org/ip2/ip2_terminology_db.cfm).
[63] Policy Cross-domain Task Force Report, 319.
[64] Preston, "General Study 09 Final Report," op. cit., 4.
[65] Ibid., 91.
[66] See, for example, case study 03 (*HorizonZero*/ZeroHorizon Online Magazine), where the creator notes that, in response to the question *What preservation strategies and/or methods are implemented and how*? (research question 19a), "[m]aintenance of the Web site will be contracted for a period of ten years, though much of the interaction (chat rooms, message boards) will be disabled" (Brent Lee (2004), "InterPARES 2 Project - Case Study 03 Final Report: *HorizonZero*/Zero Horizon Online Magazine and Media Database," 9. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs03_final_report.pdf). Likewise, in case study 02 (Performance Artist Stelarc), in response to the same research question, the creator states that "[e]ach performance generates its own documentation, each of the drawings, photos and videos made are converted to appropriate digital formats when needed, and are then posted to the Web site" (Daniel and Payne, "Case Study 02 Final Report," op. cit., 14).

the most widely held belief about digital preservation that the case studies uncovered; namely, that "it must be somebody else's problem."[67] This belief can target the specific—somebody else in the organization is in charge—or the general—for example, a state institution or some other research and resource-rich entity or organization must be doing something.

Another variation on this "it must be somebody else's problem" approach is to *purposely* transfer responsibility for long-term maintenance and preservation to somebody else. An example of this approach is provided in case study 21, where, in relation to concerns of the Supreme Court of Singapore over its ability to continue to manage the public key infrastructure (PKI) that it uses to issue digital certificates to court solicitors, the Court is considering outsourcing management of its PKI activities to a licensed certification authority so that it can be "insulated from managing technology obsolescence."[68]

Somewhat related to the "it must be somebody else's problem" pretext is a paralysis towards preservation due ostensibly to concerns about cost but that, in fact, often belie an underlying malaise or lack of will, especially among organizations, to address preservation issues and/or to provide access to preservation resources for preservation-conscious individuals within an organization's sphere of operations. Concerns regarding this latter issue were most notably expressed by scientists decrying the conspicuous lack of suitable repositories for preserving scientific data. For example, as noted earlier, there was a palpable sense of frustration (even, in some cases as the report states, desperation) expressed by many of the participants in the general study 09 (Digital Recordkeeping Practices of GIS Archaeologists) survey over the continuing lack of suitable long-term preservation repositories available for digital archaeological data and records.[69] A slightly different perspective on this situation was expressed by the creator in case study 06 (Cybercartographic Atlas of Antarctica), who noted that "[d]ata creators have few preservation incentives beyond meeting scientific and professional requirements and the peer review process,"[70] suggesting that, in many cases, individual researchers are equally culpable when it comes to the current lack of will to pursue effective, long-term preservation of scientific research records and data. As offered by the creator in case study 06, one suggestion to help address this situation is to press "[r]esearch funding agencies … to include preservation as part of their award structures and also … provide institutional support and … require policy frameworks."[71] Finally, part of the blame for the continuing reluctance of both individuals and organizations, particularly in the sciences, to acknowledge their role in the preservation lifecycle rests in the very nature of the materials requiring preservation, both in terms of their sheer quantity and their internal and external complexity. In fact, aggregations of data, and their associated documents and records, are often so large, complex (e.g., incorporating integrated multimedia, interactivity, etc.) and distributed (both within and among multiple creators), that it is difficult to find organizations that are willing, let alone able, to take on the responsibility for authentic, long-term preservation.

---

[67] A telling example of this is provided in general study 09, which characterized the rationale "as a "middle man mindset," in which the [survey] participant (especially those working as consultants) considers long-term preservation to be solely their client's problem" (Preston, "General Study 09 Final Report," op. cit., 54). As one of the study's survey participants matter-of-factly stated, "we provide the data to our clients and what they do with it is generally up to them," (Ibid.) seemingly oblivious to the fact that, in the digital environment, preservation begins with the creator and extends through to the preserver and is, therefore, a continuous and *distributed* responsibility.

[68] Goh, "Case Study 21 Final Report," op. cit., 33.

[69] Preston, "General Study 09 Final Report," op. cit., 90. See also specific participant comments regarding this issue, summarized on p. 54 in the Final Report.

[70] Lauriault and Hackett, "Case Study 06 Final Report," op. cit., 31.

[71] Ibid.

This exact situation appears to be exacerbating the preservation efforts of the creator in case study 06, who is working with the Carleton University Library "to attempt to archive the CAA [Cartographic Atlas of Antarctica], as it exists at the end of the project, as per SSHRC [Social Sciences and Humanities Research Council of Canada] requirements," since "at the moment, there are no institutions capable of archiving SSHRC supported project data and results."[72] Moreover, although it is expected that the Scientific Committee on Antarctic Research (SCAR) will eventually take over responsibility for the long-term maintenance and future development of the CAA, the creator acknowledges that this is by no means certain, given SCAR's limited human, financial and institutional resources, as well its limited technical capacity.[73]

Another variation on the publication/put-it-on-the-Web preservation strategies sees digitization as a solution. This view is generally offered either by creators: (a) who have a physical storage space problem, (b) who have trouble finding specific items among their analogue records, (c) who feel compelled (and in some cases are required by law) to oblige increasing public demand for remote electronic access to digital records and other resources[74] or (d) who see digitization as a means to effect preservation through increased access to, and diffusion of, the digitized resources.[75] For many, technology is touted as the solution to each of these concerns or issues. These assertions, generally, come from creators who are unfamiliar with the technical difficulties and the recurring costs involved in the long-term preservation of digital objects.

Where authenticity is concerned, most records creators saw little need to actively protect it in the digital environment. Reasons were varied among the individual and small group case studies and included the lack of monetary value inherent in the object;[76] the small number of practitioners in a field where everybody knew everybody and would recognize the work;[77] or the desire to make the content freely available for re-use by others, sometimes in an artistic context,[78] sometimes in a scientific context.[79] For the most part, individuals and small groups in

[72] Ibid., 18, 29.
[73] Ibid., 18.
[74] (a), (b) and (c) were all cited as motivating reasons behind the digitization of land registers discussed in case study 18 (see Blanchette et al., "Case Study 18 Final Report," op. cit.).
[75] This concern is very aptly demonstrated by the informant in case study 01 (Arbo Cyber, théâtre (?)), who, while acknowledging that the theatre "group is more concerned with digitization as a means of diffusion," nevertheless "see digitization as a better means of preservation than the traditional recordkeeping system" (Cardin, "Case Study 01 Final Report," op. cit., 48..
[76] For example, as is noted in case study 09(3), "[a]nalogue material is kept largely because of its commercial value. Individual animation cels or other artwork can be sold in galleries or at auction, but digital files have no value or only ephemeral value for such purposes. Ironically, digital files can take on value if they are touched by famous people, for example a celebrity may have written on the label of removable media even though he/she never used the computer with which the file was created" (Turner et al., "Case Study 09(3) Final Report," op. cit., 5). On the other hand, as is also stated later in the same report, "[o]ther marketing and promotional uses include re-use of digital entities for interviews with animators, party events, awards, and so on, such as value-added material on DVD versions of the studio's films" (Ibid., 8).
[77] Such as the work of performance artist Stelarc, who, for example, believes that the authenticity of the digital objects emanating from his performances "is assured primarily because of his own unique position at the centre of the entire process and the unique nature of the performance events" (Daniel and Payne, "Case Study 02 Final Report," op. cit., 12). This concept of authenticity relates closely to the concept of "personal" authenticity identified by the Domain 2 Task Force, in which "authenticity denotes the degree to which an artwork manifests the individuality and essence of its creator ... [such that] the artist is the artwork, unmediated by any records." Thus, "[t]he prevalence of this notion of authenticity explains why many artists do not concern themselves with explicitly marking the identity of their works; to them it is inconceivable that anyone else either could or would produce art like theirs. Anything an artist makes (or directs the making of) is authentic, by this definition" (Domain 2 Task Force Report, 146).
[78] The most explicit example of this is found in case study 01, where it is stated that "[t]he concept of re-use is so important to the [Arbo Cyber, théâtre (?)] group, that the informants felt that even after the dissolution of the group, artists should be able to later use records in future individual or other group projects. In this sense, the informants treat the *Ludosynthèse* not only as a testament to past performances, but also as a source of information for future use. It is creating memory" (Cardin, "Case Study 01

the arts focus seemed content with the understanding that they are the arbiters of authenticity during their lifetime. In fact, for example, in the interactive multimedia installation of case study 10, direct stewardship over the digital objects in question by their individual creators was believed by all subjects in the study to be the only reliable means of maintaining and assuring the authenticity of those objects, such that the authenticity of the objects could no longer be assured after the death of their creator.[80]

In contrast, those in organizations tended to trust that somebody else was responsible for implementing necessary measures to ensure authenticity and effect preservation—in some case studies this was true, while in others it was not. There were also examples, as mentioned earlier, of continued use of paper-based record systems, especially in environments where legal issues related to evidential value—especially in relation to intellectual property rights—might emerge at a later date.[81]

Where questions of safeguarding the identity and integrity of the records are concerned; for example, with questions of copyright and intellectual property, the findings of InterPARES 1 related to the importance of consistently and properly recorded elements of metadata was again a primary finding.[82] An interesting variation on the role of metadata was introduced in some of the arts-based case studies. As with authenticity, individual creators could only control how their works might be re-presented during their lifetime. For an accurate (and authentic) understanding of their intent for re-presentations of their works beyond that time frame, it would be necessary for the parameters of that intent to be recorded and carried forward, most probably in metadata.[83] When asked to identify, from among a list of six options, which method they believed was the best approach for representing the identity of musical works that lack a musical score (in the traditional sense), nearly one quarter (24%) of those who participated in the general study 04 survey of the recordkeeping practices of composers indicated that the identity of such works is best represented by a audio or video recording of the performance,[84] suggesting that recording a performance that fulfilled the composer's vision for the work would also provide a method to fix the composer's intent.

Perhaps the most promising and forward-thinking preservation strategy encountered by the Domain 3 researchers was observed in science focus case study 08 (NASA). Here the creator

Final Report," op. cit., 23). In fact, one of the primary purposes of the *Ludosynthèse*—the group's interactive and dynamic Web site—is to allow "spectator-users to develop or recreate performances in Arbo style using digital media" (Ibid., 3).

[79] In a sense, the Antarctic Treaty Searchable Database, which "has been designed to facilitate knowledge discovery about the policies and strategies that promote *"international cooperation*" and *the "use of Antarctica for peaceful purposes only"* [...] *"in the interest of all mankind,"* as promoted in the *Preamble* of the *1959 Antarctic Treaty*" (Berkman et al., "Case Study 12 Final Report," op. cit., 9, 34 (emphasis as in original)), also falls into this category, in as much as the database serves as a portal to public-domain policy documents whose authenticity, the database's creator insists, "can only be assured by the government agencies that issue the records" (Ibid., 29).

[80] Sally Hubbard (2006), "InterPARES 2 Project - Case Study 10 Final Report: *The Danube Exodus*," 8. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs10_final_report.pdf.

[81] See, for example, case study 03—which analyzed the recordkeeping activities of an online arts magazine (*HorizonZero*)—where it is noted that e-mail, contracts and other legal documents are kept in analogue form in a paper-based filing system apart from the creator's digital "recordkeeping" system where all the other documents and digital objects created by *Horizon*Zero are kept (Lee, "Case Study 03 Final Report," op. cit., 4, 6).

[82] See case study 06 (Cybercartographic Atlas of Antarctica) for an excellent example of this (Lauriault and Hackett, "Case Study 06 Final Report," op. cit).

[83] The necessity of such a process was clearly demonstrated in the aforementioned attempt to resurrect and stage an authentic performance of an electroacoustic work, the partial failure of which led the Domain 2 researchers to conclude that "[c]reators, while they are still living, are the best arbiters of the authenticity of performances. So it behooves them to describe their works in technologically independent (and authentically preservable) ways that will allow authentic performance in the future" (Domain 2 Task Force Report, 156).

[84] Longton, "General Study 04 Final Report," op. cit., 4 (see question 11).

claims to have implemented preservation procedures that support *persistent object preservation* (POP), which is "a technique to ensure electronic records remain accessible by making them self-describing in a way that is independent of specific hardware and software."[85] In NASA's case, this strategy relies, in part, on the use of Object Description Language (ODL)[86] to "create labels (data descriptions) for data files and other objects such as software and documents."[87] Under this strategy, technology obsolescence requires only the migration of the interpreter for the file structure description language, rather than the creation of access software for each of the file structures.[88] The fact that, in the fifteen-plus years that the PDS has been operational, "it has not been necessary to update (convert or migrate) any of the data products to other data formats,"[89] suggests that this strategy may indeed offer a viable solution for at least one aspect of the long-term preservation of digital materials.[90]

Addressing the Research Questions

Question 1

> *How do the appraisal concepts, methods and models developed by InterPARES 1 for the administrative and legal records created in databases and document management systems apply to the appraisal of the records of artistic, scientific and governmental activities resulting from the use of the technology examined by InterPARES 2?*[91]

The InterPARES 2 case studies revealed more acute examples of problems that have always challenged the preserver's ability to appraise records. Lack of metadata, idiosyncratic arrangements and fragmented storage arrangements are just a few examples. More unusual, perhaps, is the reminder that digital records, especially non-textual ones, may still have specific hardware dependencies, a situation now largely absent from the text-based recordkeeping environments of government institutions and corporations that were studied in InterPARES 1. The absence of selection criteria was also noted in relation to the types of records being studied in a number of case studies, but the development of selection criteria is not one of the preserver's

---

[85] Richard Pearce-Moses, "Persistent Object Preservation," in *A Glossary of Archival and Records Terminology* (Chicago: The Society of American Archivists, 2005). Available at http://www.archivists.org/glossary/index.asp. The use of "self-describing formats" (i.e., POP or tagging) is one of the twelve basic preservation strategies (B1.1) described in an appendix to the *Preserver Guidelines* developed by Domain 3 researchers (see Appendix 21). See also "Digital Records Maintenance and Preservation Strategies," in Appendix 21c.

[86] ODL is a language created by NASA that is used to encode data labels for its Planetary Data System (PDS) and other data systems (see National Aeronautics and Space Administration, "Chapter 12. Object Description Language Specification and Usage," in *Planetary Data System Standards Reference*, version 3.7, March 20, 2006. JPL D-7669, Part 2 (Pasadena, CA: Jet Propulsion Laboratory, California Institute of Technology, 2006). Available at http://pds.nasa.gov/documents/sr/.

[87] William Underwood (2005), "InterPARES 2 Project - Case Study 08 Final Report: Mars Global Surveyor Data Records in the Planetary Data System," 15. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs08_final_report.pdf.

[88] Ibid., 29.

[89] Ibid.

[90] It is worth noting that, based on the repeated and consistent successes of tests of the POP strategy by the Distributed Object Computation Testbed (DOCT), an interagency collaboration between the U.S. Department of Defense's Advanced Research Projects Agency, the U.S. Patent and Trademark Office and the U.S. National Archives and Records Administration (NARA), NARA regards the POP "approach as the most promising one ever suggested for preserving digital information in general, and electronic records in particular" (Kenneth Thibodeau (2001), "Building the Archives of the Future," *D-Lib Magazine* 7(2). Available at http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html), and has in fact chosen POP as the preservation method on which its soon-to-be-launched Electronic Records Archives (ERA) will rely (see http://www.archives.gov/era/about/).

[91] InterPARES 2 Project, Domain 3 Web page. Available at http://www.interpares.org/ip2/ip2_domain3.cfm.

functions seriously affected by digital technology, beyond the added importance of functionality for certain types of digital records, such as those contained in database applications.

The InterPARES 1 case studies also clearly demonstrated the need, during appraisal, to confirm whether the systems under review contained or could produce records. This situation remains essentially true for the interactive, experiential and dynamic applications studied in InterPARES 2, although the process may prove to be more complex, given the larger number of components involved and given their increasingly complex relationships.

The study of interactive, experiential and dynamic cases in InterPARES 2 considerably enriched the articulation of the nature of digital records that was provided by InterPARES 1. The implications of the findings of the InterPARES 2 case studies on the definition of a record are discussed extensively in the article *The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES*.[92] One of the major findings of InterPARES 1 was that "it is not possible to preserve a digital record: it is only possible to preserve the ability to reproduce the record."[93] That finding was based on the idea that it is not possible for a stored record written in binary code form (a form that is illegible to humans) to be capable of serving as a record. The findings of the InterPARES 2 Project, informed by its expanded scope to interactive, experiential and dynamic electronic systems, combined with a re-emphasis on the core definition of "record" and supported by precedents dating back centuries, demonstrate that this finding is valid only for a subset of digital records; namely those that, to be effective, must be manifested to a person.

InterPARES 2 case studies in the arts and in the sciences (specifically manufacturing) identified *enabling records* kept in, or associated with, computer applications or systems that are never intended to output a human-readable document. The computer instructions and specifications that enable performance of digital music works and that drive the computer-assisted manufacture of physical parts are obvious examples. Their purpose is not (re)production of a human-readable document, but (re)performance or (re)production with the mediated assistance of a computer. The definition of a record imposes no a priori constraints on the content or form of a record. The determination of whether a digital document[94] is a record depends on whether it contains and presents the necessary elements and attributes of a record, such as whether it participates in an action; has an archival bond; has an author, writer and addressee, etc.[95] Any digital document that satisfies these requirements is a record, even though it may be utterly impossible for a human to understand the record in its digital form.

Digital records are composed of one or more objects, which InterPARES 1 researchers named "digital components."[96] When "stored" (i.e., digitally encoded and placed in a storage system on digital media), these digital components necessarily have a different external form

---

[92] Duranti and Thibodeau, "The Concept of Record," op. cit., especially 46–60.

[93] Thibodeau et al., "Preservation Task Force Report," op. cit., 106.

[94] The InterPARES 2 Terminology Database defines "document" as "An indivisible unit of information constituted by a message affixed to a medium (recorded) in a stable syntactic manner. A document has fixed form and stable content," while "digital document" is defined as "A digital component, or group of digital components, that is saved [i.e., affixed to non-volatile storage on a digital medium] and is treated and managed as a document." A digital object is, in turn, defined as "A discrete aggregation of one or more bitstreams and the metadata about the properties of the object and, if applicable, methods of performing operations on the object."

[95] For a detailed discussion of all the necessary elements and attributes of a record, see the InterPARES 1 *Template for Analysis*, which is included as Appendix 1 in Duranti (ed.), *The Long-term Preservation of Authentic Electronic Records*, op. cit., 192–203. Online reprint available at http://www.interpares.org/book/interpares_book_j_app01.pdf.

[96] Defined in the InterPARES 1 Glossary as "A digital object that is part of an electronic record, or of a reproduced electronic record, or that contains one or more electronic records, or reproduced electronic records, and that requires specific methods for preservation" (http://www.interpares.org/book/interpares_book_q_gloss.pdf).

than a digital record in a form that is comprehensible to a person (i.e., as a manifested record). Further analysis in InterPARES 2 disclosed that, if kept and managed as a record, the stored digital component(s) of a manifested digital record collectively constitute a stored digital record. Thus, InterPARES 2 distinguished two sub-classes of digital records: “stored digital record,” the encoding of a digital record within a system;[97] and “manifested digital record,” a stored digital record that is visualized in a form suitable for presentation either to a person (i.e., in human-readable form) or to a computer system (i.e., in machine language). In practical terms, a stored digital record is what is kept, while a manifested digital record output from the stored digital record on a display screen or other output device is a copy that is reproduced on demand for the purpose of communicating information to persons or computer systems, rather than as a record that is kept. Thus, in such instances, a stored digital record is qualified as a record because it is intended to be used, and is capable of being used, to reproduce a manifested digital record.[98]

Analysis of the InterPARES 2 case studies showed that there is not necessarily a manifested digital record corresponding to a stored digital record. A stored digital record (or any of its digital components) may be used with—or, as in the case of enabling records, may even control—user inputs and interactions and variable data from other sources to produce changing manifestations that cannot be records precisely because their content and/or form is not fixed. Thus, a stored digital record is qualified as a record because it satisfies the definition of a record, regardless of whether it participates in the production of other records or provides a means for reproducing a manifested digital record.

InterPARES 2 refined the InterPARES 1 definition of “digital component” to one that is simpler, less ambiguous and easier to apply: “An aggregation of digitally encoded data composed of one or more bitstreams, including any metadata necessary to order, structure or manifest its content and form, requiring a given preservation action.”[99] There are three types of data that comprise stored digital documents and records: (1) content data, which are about the acts, facts or data that the document or record conveys (i.e., its content); (2) form data, which enable the system to reproduce the manifested digital document or record in the correct form from the stored digital document or record; and (3) composition data, which identify the elements of stored form and content data belonging to the document or record and map them to the different elements of structure defined by the form data. A digital component may comprise one or more of these types of data. Another important type of digital component is the digitally encoded rules that govern the reproduction of the content and form of the manifested digital document or record by determining its extrinsic elements of form, as well as any allowable content and form variations.

---

[97] Although the literal InterPARES 2 definition of stored digital record is “A stored digital document that is treated and managed as a record,” when taking into account the fact that a digital document is composed of a digital object, the effective definition becomes “A digital object, placed in a storage system on a digital medium, that is managed as a record, and which includes information about the properties of the object and may also include methods of performing operations on or with the object.” Although perhaps of more immediate relevance to the task of preservation than appraisal, it is critical to recognize that a stored digital record may not correspond to a single physical file stored on a digital medium, but rather to an amalgamation of discrete digital component(s), or subsets of components, that are logically linked to one another at the file system level through the digitally encoded information contained in a file header block, and at the application level through digitally encoded content data, form data, composition data and related rules. Thus, the digital component(s) of a stored digital record may be a subset of a stored file, a set of stored files, or various elements of data extracted from different files.

[98] In most but not all cases, enabling the reproduction of the manifested digital record is, in fact, the primary purpose of keeping the stored digital record, while the manifested record is reproduced to communicate information to persons or other systems (Duranti and Thibodeau, “The Concept of Record,” op. cit., 51).

[99] InterPARES 2 Terminology Database. Available at http://www.interpares.org/ip2/ip2_terminology_db.cfm.

InterPARES 2 analysis also proposed an addition to the established typology of records (dispositive, probative, supporting or narrative). The new type, *prospective record*, does not relate to a past action or state of affairs, but determines the form and/or content of records, actions or states subsequently created. Prospective records are subdivided into either *instructive* or *enabling* records. An instructive record is one that provides instructions, intended to be read by humans, on the creation of records in the course of some activity. It can exist either in hard copy or as a manifested digital record. Conversely, an enabling record is one that is used, in the digital form in which it is stored, by a computer system in performing an action, interaction or process and that cannot achieve its purpose if transformed into human-readable format. Thus, to be considered records, enabling records must be "properly maintained and managed [in their stored digital form] as intellectually interrelated parts of records aggregations."[100]

In this light, the findings of InterPARES 1 need to be clarified. As has already been discussed, the digital component, or set of digital components, used to reproduce an authentic copy of a human-readable record may itself be a record; namely, a stored digital record. However, it is important to understand that a stored digital record cannot be used for the purpose for which the manifested digital record is reproduced. It is at least one step removed from that purpose. In fact, many steps may be required to process the digital component(s) to output the manifested digital record in its requisite human-readable form. Thus, one might say that the purpose of the stored digital record is not to achieve the dispositive, supportive or narrative purpose of the (human-readable) record, but to enable production of authentic copies of that record. However, although that view is valid, it is overly narrow. When, for example, a student registers online to take a course, provided the university's registration system is a trustworthy recordkeeping system, the data from the student's registration record that is created from that transaction—which is kept as a stored digital record in the recordkeeping system—can be used, in combination with comparable data extracted from the stored digital registration records of other students, to produce a new record—a course roster—without going through the intermediate step either of reproducing the individual registration records in human-readable form (i.e., the manifested records), or even of reproducing manifested digital records of the individual registrations for use by the computer registration system. All that is needed is to extract the relevant data elements (i.e., course and section enrollment) from the relevant stored digital records (i.e., the registration records for each student) to produce the roster. Other data elements, such as when each student completed the course enrollment or whether the student satisfied prerequisites for the course, need not be used. Thus, the digital encodings of the enrollment records are used in a way that exactly parallels the use of paper records. When used in this manner, the *stored* digital versions of the registration records, as opposed to their *manifested* digital versions, should be recognized as the records that share an archival bond with the course roster record and should be properly managed and appraised as such.

Thus, the case studies revealed the existence of several new types of records, some with analogue equivalents and others that can exist only in a digital environment. These digital records are created within record systems with increasingly complex groups of digital components, multiplying the number of relationships that must be identified, documented, managed and preserved. There are systems that cannot display their stored digital records in human-readable form. There are also systems that can display the content of a manifested digital record in a number of documentary forms (such as spreadsheet data), although each available form is in fact fixed and stable. There are systems where non-human-readable records will

[100] Duranti and Thibodeau, "The Concept of Record," op. cit., 59.

participate in the production of digital content that never becomes fixed and is therefore not a record. And there are systems that appear to produce variable content, although careful analysis confirms that this apparently changing content is delimited by fixed rules, and therefore is, in fact, fixed and stable. Termed "bounded variability" within the computer and information sciences field, the concept describes an environment where "changes to the form and/or content of a digital record ... are limited and controlled by fixed rules [and a stable store of content data, form data and composition data], so that the same query, request or interaction always generates the same result."[101] It is important to emphasize that this is an aspect of digital records, with implications for appraisal, that goes beyond the conclusion reached in InterPARES 1 (and re-affirmed in InterPARES 2) that, although the physical integrity of a record—such as the proper number of bit strings—may be compromised, as long as the record still retains all of its essential attributes such that the message it is meant to communicate to achieve its purpose is unaltered, it can be considered "essentially complete and uncorrupted."[102]

In fact, through identification of the concept of bounded variability, InterPARES 2 has expanded the circumstances under which variations in a record's form and content may be acceptable beyond basic considerations of unintended, accidental or incidental alterations to the record's physical integrity to include consideration of a creator's intentions vis-à-vis the use of variable record presentation elements. With respect to variations in documentary form, it is important to recognize that:

> In many interactive, experiential and dynamic documents, authors or writers[57] intentionally use specific possibilities which digital technology offers for variability in the form in which information is presented. In such cases, the form is 'fixed' in that the design allows certain aspects of form to vary and not others. Documentary forms that include variable elements do not violate the requirements for fixed form, any more than analog audio and motion video recordings, which present temporal variations in sound and imagery. Such variability in presentation intended by the author should be seen as part of the extrinsic elements of the documentary form….With electronic records, then, the 'fixed' form consists of those aspects of form which the author or the writer intended or could control.[103]

Likewise, a creator may invoke this type of intentional "bounded variability" in the content of a record without compromising the record's integrity. For example, as is the case with certain online sales catalogues, interactive digital environments enable a creator to use documentary forms that permit variable selection of content and variable sequencing of that selection, such as in the display of subsets of the content of such catalogues in response to specific user input.[104] As long as the *stored digital record* is controlled by fixed rules such that the same query, request or interaction always generates the same documentary form and content selection in the *manifested digital document*,[105] such cases can satisfy the requirement for fixed content.[106]

---

[101] Excerpted from the definition for "bounded variability" from the InterPARES 2 Terminology Database, available at http://www.interpares.org/ip2/ip2_terminology_db.cfm. For a more in-depth discussion of *bounded variability*, see Duranti and Thibodeau, "The Concept of Record," op. cit., 47–49.
[102] MacNeil et al., "Appraisal Task Force Report," op. cit., 47.
[103] Duranti and Thibodeau, "The Concept of Record," op. cit., 47–48. Note: footnote reference in the quote is from the original text, and is not reproduced here.
[104] Ibid., 49.
[105] Defined in the InterPARES 2 Terminology Database as "A digital document that is visualized or rendered from a stored digital document and/or stored digital component(s) in a form suitable for presentation either to a person (i.e., in human-readable form) or to a computer system (i.e., in machine language)."

Another difference that must be taken into account during appraisals of records in interactive, experiential and dynamic digital environments is that the use of technology may change the activities of the records creator. In turn, such changes are likely to cascade into changes in records creation and recordkeeping activities. An example of the relevance of this to appraisal, and, in particular, the need to periodically re-examine a creator's document and records creation activities (and supporting technologies) after the initial appraisal, can be found in case study 24 (VanMap), where, although the creator's existing GIS system is not designed to create records, there are tentative plans to modify its architecture so that it can. This also highlights the notion that, in contrast to the emphasis in InterPARES 1 on monitoring only records that had been appraised for permanent retention, the InterPARES 2 research clearly demonstrates the need to expand the monitoring function to include those data and records that earlier appraisals decided were not worthy of preservation, as well as any systems and/or activities that earlier appraisals determined did not generate records. However, even when the records (or the systems generating them) do not themselves change, putting the records online may give rise to significant changes in their archival bond, as the ease of access often results in records being used in many more activities than previously. Given that archival appraisal should select records for preservation based on knowledge of the creator's processes and of the entire archival fonds of that creator, appraisal in the digital environment should be seen not as a one-time evaluation of a static body of records, but as an activity that needs to span the life of the records and the activities of their creators from creation to transfer to the archives for permanent preservation.[107] This, of course, greatly increases the complexity of the appraisal process, especially in relation to the more typical "one-off" appraisals of traditional analogue records.

A final difference between the process of appraisal envisioned by InterPARES 1 and the findings of InterPARES 2 relates to the stronger differentiation that now exists, in a digital environment, between retrospective records and prospective records, particularly those that play an enabling role in the (re)production of another record. Aside from the greater technical difficulty of preserving this function of reproduction, the preserver must decide whether this type of record fits within the acquisition mandate of its institution or program. Enabling records are not a completely foreign concept within past preservation practices: scores provide the basis for future performances; photographic and moving image negatives produce new prints; digital components interact to produce authentic copies of records. With respect to enabling records, however, preservation would appear to be a "prospective" commitment to the record, rather than a retrospective one.

[106] It is also important to emphasize that in cases such as this, where the manifested digital document appears to be a record, it is actually the stored digital record—which encompasses the sales catalogue's entire "palette of possibilities" (Duranti and Thibodeau, "The Concept of Record," op. cit., 32) vis-à-vis its record form and content, not just the subset presented to the user in response to a query—that is the record that is kept and used for future reference. Duranti and Thibodeau identify these types of manifested digital documents that appear to be records, but which are not actually represented internally by a stored digital record that corresponds exactly to the apparent manifested digital record, as pseudo-documents or "pseudocs" (Ibid., 27).

[107] This process is precisely encapsulated by activity A4.2.4 (Monitor Appraisal Decisions) in the Chain of Preservation model, which the model defines as "To keep track of appraisal decisions in relation to subsequent developments within the creator's and/or preserver's activities that might make it necessary to adjust or redo an appraisal, such as substantial changes to: (1) appraised records and/or their context, (2) the creator's organizational mandate and responsibilities, (3) the creator's record-making or recordkeeping activities or systems, (4) the preserver's records preservation activities or systems and/or (5) the preserver's organizational mandate and responsibilities" (Appendix 14, 224).

Questions 2 and 3

> *How do the preservation concepts, methods and models developed by InterPARES 1 for the administrative and legal records created in databases and document management systems apply to the preservation of the records of artistic, scientific and governmental activities resulting from the use of the technologies examined by InterPARES 2?*
>
> *What preservation paradigms can be applied across activities and technologies? What preservation paradigms are required for specific types of records resulting from each activity?*[108]

As was the case with the appraisal question, the case studies revealed many of the familiar problems that have always challenged the preserver's ability to safeguard records over the long term. Poor records creation practices, poor organization of the records, poor maintenance practices, an absence of long-term planning, instability of the record's physical carrier, poor storage conditions and poor documentation are just a few well-known examples from the analogue era that have been carried forward. Given the increasingly complex and, in many instances, unique nature of the digital records associated with the interactive, experiential and dynamic applications examined in InterPARES 2, it perhaps goes without saying that many of the same preservation problems that have carried over from the analogue realm are, in many instances, far more acute in the digital realm.

The InterPARES 2 research, particularly in regard to the conceptualization of records in interactive, experiential and dynamic electronic systems, has resulted in a number of paradigmatic shifts impacting digital preservation that broaden, rather than contradict, those encountered in the InterPARES 1 research. First, whereas "[o]ne of the most important findings of InterPARES 1 was the recognition and articulation of the difference between the form in which an electronic document is manifested to a person and the form in which it is stored digitally," InterPARES 2 has enriched this distinction to suggest that individual digital components, or aggregations of digital components, might themselves constitute a record or a set of records, depending on how they are instantiated in the system and how they are used by the creator.

As discussed earlier, this finding led InterPARES 2 researchers to distinguish two fundamental sub-classes of digital records; stored digital records and manifested digital records. Although a stored digital record, composed of one or more stored digital components, and the manifested digital record reproduced from those components are related, they are nevertheless distinct and, as noted earlier in the school-registration example, achieve their respective intended purposes in fundamentally different ways: the former is used in its binary code form for "presentation" to other electronic applications or systems, while the latter is translated from its stored binary code form into a form suitable for presentation to a person. This fact has important ramifications for preservation in that the specific preservation requirements for both types of records may also be distinct—informed, in part, by the way that the records are used to achieve their respective purposes. Indeed, a record, such as an enabling prospective record—or, depending on the circumstances, perhaps also a more "traditional" retrospective record—that need only be "presented" and used in binary code form to achieve its purpose will likely require

[108] InterPARES 2 Project, Domain 3 Web page. Available at http://www.interpares.org/ip2/ip2_domain3.cfm.

different preservation measures than either a retrospective or an instructive prospective record that to achieve its purpose must be reproduced in its manifested digital form for presentation to a person.

A second paradigmatic shift of the InterPARES 2 research impacting digital preservation relates, as discussed earlier, to the stronger differentiation that now exists, in a digital environment, between retrospective and prospective records, and particularly those that play an enabling role in the (re)production of another record. Enabling prospective records, in particular, offer unique preservation challenges because of the strict requirement that they be maintained in the systems in which they were created—or in systems with identical functionality—to be able to preserve their ability to produce or enable the interactions, experiences, performances or other processes they were intended to generate. Although the bitstream encoding of the stored digital components of retrospective and instructive prospective records that are kept to reproduce manifested digital records can, in many cases, be converted from one format to another (e.g., .doc to .pdf) without compromising the ability of the manifested digital record to achieve its purpose,[109] the bitstream of an enabling record must be preserved in its original form for the record to achieve its purpose.

A third paradigmatic shift relates to the concept of bounded variability and its impact on our understanding of manifested digital records. In particular, this new concept enriches the concept of the manifested record to encompass any and all types of variability of form and content that are specific to the record (i.e., represent the author's intent) and that are governed by *fixed* rules or instructions.[110]

These realizations lead, in turn, to concerns about whether this more complex and broadened range of records can be preserved according to known preservation strategies within existing preservation paradigms. As discussed earlier in the "Findings" section, the Project's case studies, which focused on records creation and maintenance, offered few concrete examples of preservation scenarios, aside from digitization and "putting it on the Web." The above discussion of variations in the form, or function, or even behaviour of digital records suggests that preservation is already possible within the known parameters of today's preservation strategies. These records clearly include more components, which exist in increasingly complex relationships. The accurate and authentic reproduction of records generated by interactive, experiential and dynamic systems will require increasingly sophisticated metadata (and/or metadata management) to document either the intent(s) of the creator or the limitations to that intent imposed by inadequacies in the available technology.[111] New types of hardware dependencies will have to be overcome. New methods to confirm the successful identification and preservation of digital records will have to be defined, especially in cases where the successful reproduction of content and documentary form in a human-readable format (i.e., manifested digital record) is no longer the goal of the preservation process.

Among other things, these issues will require that systems intended to preserve digital records be capable of the following:

1. Identifying and locating all the digital components of each record in the system.
2. Managing each digital component based on its specific preservation requirements.

[109] See Duranti and Thibodeau, "The Concept of Record," op. cit., 20.
[110] Ibid., 51.
[111] For a comprehensive analysis of the requirements and real-life context for metadata that relate to the establishment of reliability and authenticity, as well as the long-term preservation and potential re-usability of digital records, see the Description Cross-domain Task Force Report. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_6_description_task_force.pdf.

3. Identifying and managing as records those digital components that are used as records, including digital components that are records themselves (e.g., enabling records) and/or are components of one or more records.
4. Identifying and managing both the hardware and the software required for processing all digital components in the formats required for them to fulfil their function(s) as records and/or as components of stored and/or manifested digital records.
5. Applying the appropriate software, and, as required, hardware, to each component to reproduce the manifested digital record from the stored digital record and/or process the stored digital record in cases where the stored digital record is also an enabling record.

Question 4

> *What metadata are necessary to support appraisal and preservation of authentic digital records resulting from each activity?* [112]

Although it was hoped that the answer to this final Domain 3 question would flow naturally from archivists' review of the case studies, it was subsequently realized that this would not occur for two main reasons. First, as noted earlier, the focus of the case studies was on creators, nearly all of whom were neither involved in, nor interested in, appraisal of their records, and many of whom also were not involved in long-term preservation of their records. Second, despite initial hopes, it was not possible, because of the relatively small number of archivists involved in the Project, to appraise the case studies' records, nor, consequently, to determine an appropriate preservation strategy for them. Thus, very few data were collected that could be used by the Domain 3 researchers to address this question. However, the reports of both the Description Cross-domain and the Modeling Cross-domain task forces provide detailed analyses and discussions of the metadata necessary to support appraisal and preservation of authentic digital records. In fact, this final Domain 3 research question was "adopted" by the Modeling Cross-domain and integrated into its work during development of the Chain of Preservation model. Likewise, as stated in the Description Cross-domain Task Force Report, the premise underlying its work "is that detailed trustworthy metadata are key to ensuring the creation of reliable, and preservation of authentic, records and other entities in electronic systems…. [which] argues for an end-to-end metadata management regime that addresses which metadata need to be created and/or carried forward in time, for what purposes, by whom, and how they are to be preserved and validated."[113] To this end, the Description Cross-domain, together with the Modeling Cross-domain, developed a *metadata specification model*. The metadata specification model, which aligns closely with the Reference Model for an Open Archival Information System (OAIS),[114] is intended to be used, in conjunction with the Chain of Preservation model, to provide the basis for developing specifications for automated tools that can be used to systematically assist with the creation, capture, management and preservation of essential metadata for active and preserved records as well as with the identification of which metadata need to be manually created and also which can be summarized and discarded at certain points.

[112] InterPARES 2 Project, Domain 3 Web page. Available at http://www.interpares.org/ip2/ip2_domain3.cfm.
[113] Description Cross-domain Task Force Report, 291.
[114] The OAIS model describes the key roles, responsibilities and functions of a digital repository, including Ingest, "Archival" Storage, Data Management, Preservation Planning, Administration and Access. See http://public.ccsds.org/publications/archive/650x0b1.pdf.

Toward Guidelines for Preserving Authentic Digital Records

Much of the InterPARES 2 research revealed the critical absence of tools necessary to ensure the preservation of digital records. This finding was repeated across the arts, science and government focuses and affected individual creators, small groups, funded projects, government departments and large corporations. Each InterPARES research unit—the focuses, domains and cross-domains—identified missing policies, strategies, principles and guidelines to assist records creators and preservers during the ongoing transition from analogue to a digital recordkeeping environment. It fell to Domain 3 to produce the *Preserver Guidelines*,[115] a document designed to accompany the *Principles for Records Preservers*[116] developed by the Policy Cross-domain. Because the case studies focused on records creators, little was found in the way of preservation "best practices" being undertaken by the participants. Instead, the *Preserver Guidelines* reflect two perspectives:

- Actions that would have to be undertaken to avoid some of the situations encountered in the more problematic case studies
- Actions that would have to be undertaken to address the appraisal and preservation concerns identified in the InterPARES 2 research

[115] See Appendix 21.
[116] See the *Policy Framework* in Appendix 19.

PART FIVE

MODELING DIGITAL RECORDS CREATION, MAINTENANCE AND PRESERVATION

Modeling Cross-domain Task Force Report

Chain of Preservation Model Narrative

by
Terry Eastwood, The University of British Columbia
Randy Preston, The InterPARES Project

Business-driven Recordkeeping Model Narrative

by
Hans Hofman, National Archives of the Netherlands

Preamble

The objective of the InterPARES 2 Project is to develop a theoretical understanding of the records generated by experiential, interactive and dynamic systems, of their process of creation, and of their present and potential use in the artistic, scientific and governmental sectors.[1] The Modeling Cross-domain Task Force, while not included in the original Project plan, was added to three other cross-domain task forces for two reasons:

1. Modeling of the creation, maintenance and preservation of records, both from the preserver viewpoint and from the business process viewpoint, presents a picture of all the activities of records making, records keeping, and records preservation and their relationship for researchers in the Project to utilize.
2. Modeling serves as a method of analysis to which researchers in the other domains and focuses could contribute their expertise to develop better understanding of requirements for long-term preservation of records in experiential, interactive and dynamic systems.

No specific research questions were formulated for the Modeling Cross-domain at the outset of the Project. However, some of the questions of Domain 3 were addressed by the Modeling Cross-domain. These are as follows:

- What preservation paradigms can be applied across activities and technologies? What preservation paradigms are required for specific types of records resulting from each activity?
- What metadata are necessary to support appraisal and preservation of authentic digital records resulting from each activity?[2]

The modeling work was based upon previous projects such as the UBC Project[3] and the InterPARES 1 Project,[4] as well as upon theoretical concepts developed through research undertaken by the Records Continuum Research Group in Australia.[5]

The Modeling Cross-domain developed two models. The Chain of Preservation (COP) Model is based on the perspective of the entity responsible for long-term preservation of digital records. The Business-Driven Recordkeeping (BDR) Model is based on the perspective of the records creating entity. The two models are mutually supportive in that they provide two ways of looking at the problem of long-term preservation of authentic digital records. The COP Model is explained in the next part of this report, followed by an explanation of the BDR model. The purpose of developing two models was to be able to look at the question of preservation of the

[1] Luciana Duranti (2004), "InterPARES 2 Project Midterm Report to the Social Sciences and Humanities Research Council." MCRI Grant No. 412-2001-1003, 1 (unpublished).
[2] Luciana Duranti (2001), "International Research on Permanent Authentic Records in Electronic Systems (InterPARES): Experiential, Interactive and Dynamic Records," SSHRC MCRI InterPARES 2 Project Proposal, 412-2001, 14. Available at http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf.
[3] See Luciana Duranti and Heather MacNeil (1996), "The Protection of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project," *Archivaria* 42 (Fall): 46–67. Online reprint available at http://journals.sfu.ca/archivar/index.php/archivaria/article/view/12153/13158; and Luciana Duranti, Terry Eastwood and Heather MacNeil, *Preservation of the Integrity of Electronic Records* (Dordrecht, Kluwer Academic Publishing, 2002). An online reprint of the book is available at http://www.interpares.org/book/index.cfm. See also the UBC Project's Web site at http://www.interpares.org/UBCProject/index.htm.
[4] International Research on Permanent Authentic Records in Electronic Systems (InterPARES 1 Project). The Project's Web site is available at http://www.interpares.org/ip1/ip1_index.cfm.
[5] See the Web site of the Records Continuum Research Group, Monash University. Available at http://www.sims.monash.edu.au/research/rcrg/index.html.

reliability, authenticity and accuracy of records in experiential, interactive and dynamic systems through the lens of two somewhat different conceptions, each offering its own insights.

The result of the modeling work undertaken within the Modeling Cross-domain, as presented in this narrative, constitutes an intermediate version of the models. Both models, therefore, should be considered *consultation drafts*. While they have been discussed extensively within the Project, further discussion from other viewpoints and with other experts is still required so that these models can benefit from such dialogue and evolve to a more refined stage.

Further work is needed on the definitions, the description of the diagrams and the validation of the models. The COP model has, for instance, not been validated in the artistic and governmental sectors, while the BDR model has not been validated in the artistic and scientific sectors and also needs more testing within administrative environments by doing walkthroughs based on the available and new case studies.[6] The current versions, however, have been sufficiently developed to provide insight into the approach taken and an understanding of the underlying concepts. Both models are based on experience and, as such, on a logical construction of existing knowledge. They are now offered as instruments for feedback from individuals and organizations.

Chain of Preservation Model

Introduction

Early in the work of the InterPARES 2 Project, which began in 2001, the Modeling Cross-domain Task Force began developing a model that would eventually be called the Chain of Preservation (COP) model.[7] The COP model, which depicts all the activities and the inputs and outputs that are needed to create, manage and preserve reliable and authentic digital records, was created based on the IDEF0 (Integrated Definition Function) modeling process using IDEF0 modeling software. IDEF0 is a U.S. Federal Information Processing Standard for function modeling.[8] A function model is a structured representation of the functions, activities and processes within the modeled system or subject area.[9] The COP model consists of a series of diagrams depicting all the activities involved in the life-cycle management of digital records together with a glossary of all the terms appearing on the diagrams. The diagrams and glossary are included in Appendices 1 and 2 and are also available on the InterPARES 2 Web site.[10]

The COP model was based on three previous models. The first of these previous models was generated during a project entitled "The Preservation of the Integrity of Electronic Records"

[6] Two walkthroughs of an earlier draft of the COP model were conducted using data from two scientific sector (Focus 2) case studies. See William Underwood, Kevin Glick and Mark Wolfe (2007), "InterPARES 2 Project - General Study 12 Final Report: Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs12_final_report.pdf; and Randy Preston (2004), "InterPARES 2 Project - Modeling Cross-domain: Walkthrough of the Manage Chain of Preservation Model Using Case Study 14 Data," draft report. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs14_COP_model_walkthrough.pdf.

[7] The COP model was also called the Grand Unified Model (GUM), Manage Electronic Records (MER) model, Manage Records Lifecycle (MRL) model and the Manage Chain of Preservation (MCP) model during the earlier stages of its development.

[8] See National Institute of Standards and Technology, *Draft Integration Definition for Function Modeling (IDEF0), Federal Information Processing Standards Publication 183* (Gaithersburg, MD: Computer Systems Laboratory, National Institute of Standards and Technology, 1993). Available at http://www.itl.nist.gov/fipspubs/idef02.doc.

[9] For more information about the IDEF0 process, see http://www.idef.com.

[10] See http://www.interpares.org/ip2/ip2_models.cfm. Enter terms used in the model into the search box provided to see their definitions.

(commonly referred to as the UBC Project) conducted between 1994 and 1997 by Professors Luciana Duranti and Terry Eastwood and research assistant Heather MacNeil of the University of British Columbia. Its purpose was to define the activities for the genesis and preservation of an agency's archival fonds. Its scope was to control records (archival documents) according to the agency's mandate using the principles of archival science. It adopted the perspective or viewpoint of the records creator. The diagrams and glossary constituting the UBC Project's model are available on the InterPARES Web site[11] and in the book publishing the results of the Project.[12] The other two previous models were produced also using IDEF0 methodology during 1998-2001 by the InterPARES 1 Project; one of the selection function and the other of the preservation function.[13] The "Select Electronic Records" model defines all the activities involved in the selection of authentic digital records for long-term preservation. It adopts the viewpoint of the entity responsible for long-term preservation of digital records. It covers all the activities of the preserver in appraising and carrying out the disposition of digital records. The "Preserve Electronic Records" model takes the same viewpoint and covers all the activities conducted in preserving authentic digital records. These two models are also available on the InterPARES Web site[14] and in the book publishing the results of the InterPARES 1 Project.[15]

The process of developing the COP model involved examining the three pre-existing models and their conceptual bases closely, and then engaging in producing a new model based on previous conceptions but refining them and creating new definitions of terms where required. The work was done at the University of British Columbia with the assistance mainly of research assistants Rachel Mills and Bart Ballaux and, later in the work, Randy Preston. Luciana Duranti, the Project Director, participated in the early stages of the work. As the modeling exercise progressed, the interim results were presented to the wider research group for their criticism and comment, which often lead to improvements in the model.

Purpose, scope and perspective of the model

In one sense, the purpose of the COP model was to unify the three previously produced models, which each depicted one of the functions in the life-cycle management of digital records. As is widely recognized, digital records must be carefully managed throughout their entire existence to ensure that they are accessible and readable over time with their form, content and relationships intact to the extent necessary for their continuing trustworthiness as records. It is also widely recognized that management of digital records must proceed from a comprehensive understanding of all phases or stages in the lifecycle of records, from the time they are generated, through their maintenance by their creator, and during their appraisal, disposition and long-term preservation as authentic memorials of the actions and matters of which they are a part. The COP model has, then, within its scope all these phases or stages in the lifecycle of digital records and all the activities and important, specific actions that must be undertaken to ensure that digital records are properly generated in the first instance, maintain their integrity over time, and can be

[11] See http://www.interpares.org/UBCProject/index.htm.
[12] Duranti et al., *Preservation of the Integrity of Electronic Records*, op. cit.
[13] It is worth noting that the Open Archive Information System (OAIS) reference model, now an international standard (ISO 14721:2003), served as the basis for the preservation function model from InterPARES 1 and thus can, by extension, be said to have also informed the COP model.
[14] See http://www.interpares.org/book/index.cfm.
[15] Luciana Duranti (ed.), *The Long-term Preservation of Authentic Electronic Records: The Findings of the InterPARES Project* (San Miniato, Italy: Archilab, 2005). More specifically, Appendix 4: A Model of the Selection Function and Appendix 5: A Model of the Preservation Function, 239–292.

authentically reproduced at any time throughout their existence. As well, it characterizes the data and information that must be gathered, stored and utilized during the various processes of management throughout the lifecycle. The COP model also depicts the constraints or controls on the various activities and actions it characterizes.

The matter of the viewpoint or perspective of the COP model is vital to its understanding. The UBC Project model was produced from the perspective of the records creator; that is, any entity with responsibility for long-term preservation of digital records. It could be an entire organization, an agency within an organization, or even an office within an organization or agency if such decentralization of records management were appropriate. In contrast, the two InterPARES 1 Project models took the perspective of the records preserver; that is, the entity responsible for long-term or continuing or enduring preservation of authentic digital records. It could be a separate archival institution, such as a public archives responsible for preservation of a government's records, or an archival program or unit within an organization, such as a church, business or university archives, or even a repository preserving archives of records creators with which it has no organizational or administrative connection. Obviously, however, this last example presents special problems for long-term preservation of digital records. In fact, the measures taken to control digital records creation and maintenance in a setting where there is an established relationship between records creators and a designated records preserver[16] are much more difficult to articulate when the preserver has no organizational relationship with the records creators whose records it aims to preserve over the long term.

In any event, developers of the COP model faced a serious problem in unifying or rationalizing the three pre-existing models, for they did not proceed from the same perspective. This is precisely the problem that has generated criticism of the lifecycle concept, which all too often finds it difficult to rationalize or coordinate all the activities of records creation, maintenance and use by the records creator, with appraisal, disposition and long-term preservation by the records preserver. The developers of the COP model solved this problem, in effect, by adopting the perspective of *the needs of the records*. It is quite obvious that early phases or stages in the records lifecycle are managed by the records creator and later phases or stages by the records preserver, but the COP model proceeds from an understanding of the theory, methods and practice of proper records processes throughout the lifecycle of the records. This understanding recognizes that whoever takes responsibility for undertaking actions on digital records must do so with the whole sweep of the records' existence in mind. Nevertheless, because the goal of the InterPARES Project is to develop knowledge of the requirements for long-term preservation of authentic digital records, the COP model, it can be conceded, views all the activities it depicts from the perspective of that goal of preservation of authentic digital records, which, of course, is consonant with the perspective of the preserver.

Finally, it should be noted that the phrase "chain of preservation" was consciously chosen as the title for the model to indicate that, from the perspective of long-term or continuing or enduring preservation, all the activities to manage records throughout their existence are linked, as in a chain, and interdependent. If a link in the chain fails, the chain cannot do its job. If certain activities and actions are not undertaken on records, their trustworthiness (that is, their reliability and authenticity) and preservation are imperilled.

[16] A *designated records preserver* is the entity responsible for taking physical and legal custody of, and preserving (i.e., protecting and ensuring continuous access to) authentic copies of a creator's inactive records. The role of the designated records preserver should be that of a *trusted custodian* for a creator's records; that is, a preserver who can demonstrate that it has no reason to alter the preserved records or allow others to alter them and is capable of implementing all of the requirements for the authentic preservation of records.

Overview of the Chain of Preservation Model[17]

Manage Chain of Preservation (A-0)

This top-level diagram delineates the subject of the model and its overall context represented by a single box with its bounding arrows. The bounding arrows—representing the model's primary inputs, outputs, controls/constraints and mechanisms—interface with functions outside the subject area being modeled, thus establishing the model's focus.

Constraints on records activities

The IDEF0 methodology requires that constraints or controls on functions be identified. Records are always generated and kept with certain constraints in mind. The first constraint on the model has already been mentioned. It proceeds from an understanding of *archival science*; that is to say, the model is founded on archival concepts, methods and practices, which were articulated in the models produced prior to this exercise of producing the COP model.

The other constraints all relate to the context in which records come into being and exist. Records creators and preservers operate in this context. The first aspect is the context of the *juridical system*. The InterPARES Project adopts from diplomatics a definition of the term juridical system. It is a social group that is organized on the basis of a system of rules and that includes three components: the social group, the organizational principle of the social group and the system of binding rules recognized by the social group. The records produced in a society are, to a considerable extent, influenced by its social setting, by the kind of a society it is, by the way it governs itself and, in particular, by the rules and customs by which it abides. Thus, aspects of the juridical system constrain or influence the way records are made, kept, communicated, used, maintained and preserved. Laws and regulations most directly affect records processes. There may be laws or regulations that have to be observed when generating or managing records or when maintaining and preserving them. For instance, governments pass laws or promulgate regulations affecting their own records and the records of private bodies. Managing the chain of preservation has to take such rules into account.

In any given context, records may be generated and treated according to various *international, national or other standards*. Standards lay out rules or guidelines to regulate activities, practices and the like. In any given context, it may be either necessary or desirable to adhere to standards. These standards act as a further constraint on managing digital records.

The degree to which digital technology is available and the capabilities it has constitute a further constraint. In any real situation, the *state of technology* may severely limit what can be done in the management of digital records. The COP model was constructed with an understanding of the capacity of information technology as it now exists. So, for instance, the model assumes that software for electronic document and records management exists and can be utilized and that technological means exist to migrate records. In other words, it assumes an ideal situation, not one in which fiscal or other limitations prevail.

The final constraint, the *preserver's mission*, is the one that most clearly indicates the perspective of the model. The preserver's mission is identified as a constraint because the main interest animating the model is the long-term preservation of reliable, accurate and authentic digital records. Now, obviously, there are additional constraints on record-making and recordkeeping that are dictated by the mission, mandate and administrative habits of the records

[17] The COP model IDEF0 diagrams, together with the activity and arrow definitions, are provided in Appendix 14.

creator. Therefore, the mission, mandate and functions of the records creator act as a constraint on its management of its digital records. This constraint was recognized in the UBC model, and the COP model was created with this recognition in mind.

Mechanisms instrumental for record activities

Many resources are needed to generate, maintain, control and preserve digital records. The COP model assumes that both the records creator and the records preserver will assign personnel tasks for records creation and management, and records preservation, respectively. Various tools, such as information technology and other equipment and supplies, will be needed to manage the lifecycle of records, as will physical facilities and infrastructure. In a real situation, lack of these resources obviously constrain what can be done, and often seriously. In the case of the model, a situation of optimum resources is assumed to allow users to picture the ideal towards which, in any given real situation, they may work.

Inputs to record activities

By definition, all of the top-level inputs to the model represent information or objects that originate outside of the main activity being modeled. The first two inputs, *creator's existing records* and *information about the creator's existing records*, acknowledge the fact that, aside from rare instances where the model is used to support the development of a chain of preservation framework for a new records creating entity, any entity to which the model is being applied will already have in place some level of record-making and recordkeeping capacity, however modest. In other words, it is assumed that most creators will have already created records prior to developing and implementing the integrated chain of preservation framework envisioned by this model. Thus, first two inputs represent records made or received by the creator that predate the development and implementation of the chain of preservation framework and that will need to be incorporated into the creator's "new" record-making and recordkeeping systems following implementation of the framework.

The next input, *received documents*, represents inputs received by the creator from external juridical or physical persons subsequent to implementation of the framework. In some cases, the creator may require additional *information about the received documents' context* to properly *identify* (i.e., attach to the documents identity metadata that convey the action in which the documents participate and their immediate contexts—see A2.2.3), *declare* (i.e., intellectually set the documents aside as records by registering and classifying them—see A2.2.4), and/or *execute* (i.e., attach to the declared records metadata that convey information about the formal execution phase of the administrative procedure in which the record participates—see A2.2.5) them. Although the model does not specify exactly how, where or from what source this supplementary information is or should be obtained, it is assumed that the information would either be generated and supplied by the external entity transmitting the document to the creator in response to at request from the creator, or be compiled by the creator through examination of other information available about the external entity and/or the action to which the received documents relate.

In addition to analysis of the creator's existing records (and information about the records), *information about the creator* and *information about the (designated) preserver*—such as the creator's and preserver's mission, organizational structure, activities and existing technological, financial and human resources, etc.—are needed to help identify the framework policies and design requirements for the chain of preservation framework (see A1.1).

Information about available technology (i.e., documentation concerning the software and hardware available on the market to the creator and to the preserver) is required to support the (re)design for the record-making (see A1.2.1), recordkeeping (see A1.2.2) and permanent preservation systems (see A1.2.3). It is also needed during the appraisal of the creator's records by the preserver to support the process of reconciling the preservation requirements of the records being appraised with the preserver's existing and/or potential preservation capabilities (see A4.2.2.3.3).

Ideally, creator's and preservers should, to the extent possible, establish and provide users with access to formal records discovery instruments (indexes, inventories, record descriptions, etc.) as well as put in place formal records access submission procedures. The final two top-level inputs, *unmediated access requests for kept records and/or information* and *unmediated access requests for preserved records and/or information*, recognize that this level of control is not always feasible.

Outputs of record activities

Although there are many different top-level outputs, closer inspection reveals that they are all either documents or records[18] and could, in fact, be represented by just two arrows: *documents issued by the creator* and *documents issued by the preserver*. The primary reason choosing instead to identify the various types of documents and records generated by the creator and the preserver at this level is to provide users of the model with an high-level overview of the main types of record-making, recordkeeping and records preservation activities that generate output that needs to be managed.

The Four Main Records Activities

The COP model distinguishes four main records activities: (1) managing the framework for the chain of preservation, (2) managing records creation, (3) managing records in a recordkeeping system and (4) preserving selected records. The following section summarizes these main activities and their important outcomes.

Managing the Framework for the Chain of Preservation (A1)

This activity involves determining framework requirements and designing, implementing and maintaining a chain of preservation framework. As used in this context, a *framework* means all the elements of policy, strategy, method and so on needed to manage digital records.

Developing the Management Framework (A1.1)

This activity involves analyzing information about the records creator and its existing records and about the designated records preserver to identify the management policies and design requirements for the chain of preservation framework. Specifically, the component activities of developing the management framework are:

1. *to analyze the records creator* (A1.1.1), which involves assessing key information about the records creator to help identify the record-making and recordkeeping-related requirements for the chain of preservation framework, including the creator's:
    - mission;
    - organizational structure;

[18] Technically speaking, given the definition of record used by the InterPARES Project, all of the outputs are in fact documents even though many are identified as records. This is because, by definition, a record is a document that an entity either makes or receives and *sets aside*. From the perspective of the sender, things that send to external juridical or physical persons are documents, while the copies of the sent documents retained by the sender are records.

- activities;
- existing technological, financial and human resources; and
- records-related needs and risks.

2. *to analyze the creator's existing records* (A1.1.2), which involves assessing creator's existing records and information about those records to determine framework requirements for managing the records;
3. *to analyze the designated records preserver* (A1.1.3), which involves assessing the same key information about the designated records preserver as noted above for the records creator to help identify the preservation-related requirements for the chain of preservation framework;
4. *to establish the framework management policies* (A1.1.4), which involves developing management regime policies for establishing overall framework design requirements based on the findings of the above analyses of the designated preserver and the records creator and its existing records; and
5. *to establish the framework design requirements* (A1.1.5), which involves first identifying the rules guiding the design of the chain of preservation framework on the basis of the analyses of the designated preserver and the records creator and its existing records, and then issuing specific design requirements—that is, the record-making, recordkeeping and permanent preservation needs that should guide the framework design—and framework policies—that is, the collective, high-level management principles that should guide and control the development of the framework requirements.

Designing the Framework (A1.2)

Designing the chain of preservation framework involves developing three things: (1) a record-making system design, (2) a recordkeeping system design and (3) a permanent preservation system design.

The distinction between the record-making and recordkeeping systems should be made clear. The record-making system consists of a set of rules governing the creation of records and the tools and mechanisms used to implement these rules. In practice, this involves the *capture* and *identification* of documents, the *declaration* and *execution* of records and the deliberate act of *setting aside* completed records by transferring them to the recordkeeping system. Thus, the record-making system can be distinguished conceptually from the recordkeeping system, which is a set of rules governing the controlled *access* to, and the *storage*, *maintenance* and *disposition* of, *kept records* (i.e., completed records that have been "set aside" in the recordkeeping system for future action or reference) and/or information about kept records and the tools and mechanisms used to implement these rules. In real situations, this distinction might be transparent to the user.

Designing the Record-making System (A1.2.1), the Recordkeeping System (A1.2.2) and the Permanent Preservation System (A1.2.3)

The process for designing the record-making, recordkeeping and permanent preservation systems involves the same four basic activities for each system (as discussed below). The major design distinctions between the three systems relate primarily to the types of "sub-systems" within each system and, to varying degrees, the types of rules and procedures and system instruments associated with each system. As envisioned in this model, the so-called "sub-systems" within each of the three main systems are distinguished primarily on the basis of conceptual rather than concrete functional boundaries relating to the main types of activities that impact upon the creation and maintenance of accurate and reliable records and the long-term preservation of authentic records. Thus, as with the distinction between the record-making and recordkeeping systems, the distinctions between some or all of the sub-systems within each of the three main systems may, in many cases, be transparent to the user. With this understanding in mind, the key sub-systems for each of the three main systems as delineated as follows:

1. Record-making System:
   a. *Document Capture System*, the main purposes of which are to facilitate the recording and saving of particular instantiations of incoming external documents or internal documents made by the creator in accordance with the specifications of the creator's documentary forms, integrated business and documentary procedures and record-making access privileges.
   b. *Document Identification System*, the main purposes of which are to facilitate the creation and recording of identity metadata for captured documents that convey the action in which the document participates and its immediate context.
   c. *Records Declaration System*, the main purpose of which is to facilitate the intellectual "setting-aside" of captured and identified documents as records by assigning classification codes from the classification scheme, and adding these codes to the record's identifying metadata, and by assigning registration numbers based on the registration scheme, and adding these numbers to the record's identifying metadata.
   d. *Records Execution System*, the main purposes of which are to facilitate the creation and recording of the key metadata (e.g., priority of transmission, transmission date, actions taken, etc.) for each record that convey information related to, and actions taken during the course of, the formal execution phase of the administrative procedure in which each record participates.
   e. *Records Transfer System*, the main purpose of which is to facilitate the transfer of executed or completed records deemed worthy of retention by the creator to the recordkeeping system for the purpose of maintaining them for future action or reference.
2. Recordkeeping System:
   a. *Records Information System*, the main purposes of which are to facilitate the generation, capture, compilation and management of information about records in the recordkeeping system, and about all recordkeeping maintenance, access and disposition activities applied to the records, for use in maintaining the authenticity of the records in the recordkeeping system and for facilitating records appraisal activities by the preserver and records indexing, storage, access and disposition activities by the creator.
   b. *Records Indexing System*, the main purpose of which is to facilitate the indexing of records in the recordkeeping system, through assignment of access points to each record using a controlled recordkeeping vocabulary, for the purpose of facilitating effective and efficient discovery and retrieval of records in the recordkeeping system.
   c. *Records Storage System*, the main purposes of which are to facilitate the processes of placing and maintaining the digital components of the creator's records, and their metadata, in a storage system on digital media.
   d. *Records Retrieval System*, the main purpose of which is to facilitate the extraction of copies of the digital components of the creator's records in storage, and their metadata, in response to retrieval requests.

e. *Records Access System*, the main purpose of which is to facilitate access to authentic copies of the creator's records, or information about the records, to authorized users upon request by connecting users with the necessary tools (such as indexes) for locating records and by effecting the reconstitution of retrieved digital components and/or information in authentic form and the presentation of the manifested records or information to users, and/or the packaging of retrieved digital components with information about how to reconstitute and present the records and/or information with the appropriate extrinsic form.
f. *Records Disposition System*, the main purposes of which are to facilitate the authorized destruction and/or transfer to a designated preserver of the creator's records in accordance with the creator's retention schedule and the preserver's terms and conditions of transfer and disposition rules and procedures.

3. Permanent Preservation System:
   a. *Records Information System*, the main purposes of which are to facilitate the generation, capture, compilation and management of information about records in the permanent preservation system, and about all preservation selection, acquisition, description, storage, retrieval and access activities applied to the records, for use in maintaining authentic copies of the creator's records in the permanent preservation system and for facilitating all records preservation activities.
   b. *Records Selection System*, the main purposes of which are to facilitate the appraisal and selection of the records of creators, as well as the ongoing monitoring of appraised records, changes to the record-making and recordkeeping activities of the creators whose records have been appraised and any other factors that might significantly impact on the accuracy, validity or integrity of the preserver's current or active appraisal decisions.
   c. *Records Acquisition System*, the main purposes of which are to facilitate the transfer, intake, processing and accessioning of the records of creators selected for long-term preservation.
   d. *Records Description System*, the main purposes of which are to facilitate the intellectual and physical control of accessioned records by recording information about their nature, make-up and contexts (juridical-administrative, provenancial, procedural, documentary and technological),[19] as well as about any changes the records have undergone since they were first created.
   e. *Records Storage System*, the main purposes of which are to facilitate the processes of placing and maintaining the digital components of the accessioned records, and their metadata, in a storage system on digital media.
   f. *Records Retrieval System*, the main purpose of which is to facilitate the extraction of copies of the digital components of the preserved records in storage, and their metadata, in response to retrieval requests.
   g. *Records Access System*, the main purpose of which is to facilitate access to authentic copies of the preserved records, or information about the records, to authorized users upon request by connecting users with the necessary tools (such as record descriptions and complimentary description instruments such as guides, inventories, indexes, repository locators and related finding aids) for locating records and by effecting the

[19] These contexts are defined in the report of the Authenticity Task Force (2001), "Appendix 1: Template for Analysis," in Duranti, *Long-term Preservation* , op. cit., 198–199. Online reprint available at http://www.interpares.org/book/interpares_book_j_app01.pdf.

reconstitution of retrieved digital components and/or information in authentic form and the presentation of the manifested records or information to users, and/or the packaging of retrieved digital components with information about how to reconstitute and present the records and/or information with the appropriate extrinsic form.

As noted above, the process for designing the record-making, recordkeeping and permanent preservation systems involves the same four basic activities for each system. These activities are identified as follows:

6. *to develop the system's administrative infrastructure* (A1.2.1.1, A1.2.2.1 and A1.2.3.1, respectively), which for each system (and sub-system) involves defining, analyzing, creating and documenting a comprehensive, integrated set of administrative policies, strategies, rules and procedures, and instruments to support system activities and to enable the systems to meet their functional requirements. *Policies* are the collective, high-level management principles that guide and control development, implementation and execution of a system/sub-system and the activities it supports. *Strategies*, which help enable policies, are authoritative objectives and methods governing the operation of a system/sub-system and the activities it supports. *Rules and procedures*, which, in turn, help enable strategies, are the authoritative instructions governing the operation of a system/sub-system and the activities it supports. Finally, within the context of this model, *instruments* are the tools that support the various record-making, recordkeeping and record preservation processes in each of the three main systems and each of their sub-systems. The key rules and procedures and system instruments applicable to each of the three main systems are described below.
   - *Rules and Procedures:*
     Within the record-making system, aside from the rules and procedures specific to the activities within each sub-system, the general rules and procedures include (see A1.2.1.1.3):
     a. *integrated business and documentary forms*, consisting of procedures for carrying out the creator's business that have been linked to a scheme or plan for organization of the creator's records;
     b. *procedures for ensuring the accuracy of records*, consisting of authoritative procedural orders designed to ensure that records are created accurate in the record-making system;
     c. *procedures for ensuring the reliability of records*, consisting of authoritative procedural orders designed to ensure that records are created reliable in the record-making system; and
     d. *record-making access privileges*, consisting of specification of the authority to compile, annotate, read, retrieve, transfer and/or destroy records in the record-making system, granted to officers and employees of the creator;

     Within the recordkeeping system, aside from the rules and procedures specific to the activities within each sub-system, the general rules and procedures include (see A1.2.2.1.4):
     a. *procedures for maintaining authentic records*, consisting of authoritative procedural orders designed to ensure that records maintain their identity and integrity as they are managed and maintained in the recordkeeping system; and
     b. *recordkeeping access privileges*, consisting of specification of the authority to annotate, read, retrieve, transfer and/or destroy records in the recordkeeping system, granted to officers and employees of the creator;

Within the permanent preservation system, aside from the rules and procedures specific to the activities within each sub-system, the general rules and procedures include (see A1.2.3.1.4):

a. *procedures for assessing the authenticity of records*, consisting of authoritative procedural orders designed to facilitate evaluation of the authenticity of the creator's records during appraisal and/or acquisition of the records by the preserver;
b. *procedures for maintaining authentic copies of records*, consisting of authoritative procedural orders outlining pre-established requirements for maintaining authentic copies of the creator's records in the custody of the preserver; and
c. *preservation access privileges*, consisting of specification of the authority to compile, annotate, read, retrieve, transfer and/or destroy records in the permanent preservation system, granted to officers and employees of the creator;

- *System Instruments*:

Within the record-making system, the key system instruments include (see A1.2.1.1.4):

a. *records forms*, consisting of specifications of the documentary forms for the various types of records of the creator;
b. *record-making metadata schemes*, consisting of lists of all necessary record-making metadata to be recorded to ensure the reliability, accuracy, identification and integrity of records created in the record-making system;
c. *record-making reporting schemes*, consisting of plans for the systematic generation of documentation or reports of the creator's record-making activities according to logically structured conventions, methods and procedural rules; and
d. *record profile schemes*, consisting of plans for the systematic generation of digital forms designed to contain the attributes of records that attest to their identity and integrity, and which are generated when users create, send and/or close records, are updated when users subsequently modify or annotate completed records, and remain inextricably linked to the records for the entire period of their existence while in the custody of the creator;

Within the recordkeeping system, the key system instruments include (see A1.2.2.1.4):

a. *recordkeeping metadata schemes*, consisting of lists of all necessary recordkeeping metadata to be recorded to ensure the identification and integrity of records maintained in the recordkeeping system;
b. *recordkeeping registration and classification schemes*, consisting of a method for assigning a unique registration number to each record in the recordkeeping system and a plan for the systematic identification and arrangement of business activities and records into categories according to logically-structured conventions, methods and procedural rules;
c. *recordkeeping reporting schemes*, consisting of plans for the systematic generation of documentation or reports of the creator's recordkeeping activities according to logically structured conventions, methods and procedural rules; and
d. *retention schedule*, consisting of a document providing description of records series and/or classes and specifying their authorized dispositions;
e. *controlled vocabulary/thesaurus*, consisting of one or more managed sets of purposefully delimited and standardised terms, phrases and concepts used by the creator to control the values of metadata elements;

Within the permanent preservation system, the key system instruments include (see A1.2.3.1.4):

a. *preservation metadata schemes*, consisting of lists of all necessary preservation metadata to be recorded to ensure the identification and integrity of records preserved in the permanent preservation system;
b. *records transfer registration and accession schemes*, consisting of a method for assigning a unique identifier to each received and accessioned records transfer, respectively;
c. *preservation reporting schemes*, consisting of plans for the systematic generation of documentation or reports of the preserver's preservation activities according to logically structured conventions, methods and procedural rules; and
d. *controlled vocabulary/thesaurus*, consisting of one or more managed sets of purposefully delimited and standardised terms, phrases and concepts used by the preserver to control the values of metadata elements;

7. *to establish the system's functional requirements* (A1.2.1.2, A1.2.2.2 and A1.2.3.2, respectively), which for each system (and sub-system) involves developing and documenting comprehensive and integrated performance, monitoring and technological system requirements. *Performance requirements* are the operational and administrative specifications for measuring the continuing ability of a system/sub-system to fulfil its purpose. *Monitoring requirements* are the operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of a system/sub-system in relation to the established performance requirements for the system/sub-system. *Technological requirements* are the specifications of the hardware and software needed for a system/sub-system.
8. *to establish the system's performance requirements* (A1.2.1.3, A1.2.2.3 and A1.2.3.3, respectively), which for each system (and sub-system) involves developing operational benchmarks or standards for the operation of the system, in relation to the established functional requirements, against which the continuing performance and adequacy of an activity, function, process, sub-system or structure within the system can be measured; and
9. *to design the system's functional infrastructure* (A1.2.1.4, A1.2.2.4 and A1.2.3.4, respectively), which for each system (and sub-system) involves developing a comprehensive, integrated design for the system and each of its sub-systems.

Implementing the Framework (A1.3)

This activity involves acquiring, testing and activating all the components of the record-making, recordkeeping and permanent preservation systems and issuing information about any problems encountered during implementation of any of these systems so that the framework design can be modified accordingly.

Maintaining the Framework (A1.4)

This ongoing activity involves: (a) assessing information about the performance of the record-making, recordkeeping and permanent preservation systems, based on analysis of performance reports periodically generated by the activities in each system and on updated information about the creator and preserver, (b) as appropriate, making recommendations about revising the framework design to correct problems or to accommodate newly-identified requirements and (c) periodically generating requests for updated information about any significant changes to the juridical-administrative, technological and related contexts within which the records creator carries out its records-related activities.

Managing Records in a Record-making System (A2)

This second main activity involves overseeing and coordinating all the activities associated with managing the making and receipt of digital records, transferring created records to the recordkeeping system (i.e., "setting aside") and monitoring the overall operation of the record-making system. It is understood that documents will be generated in the course of the activities of the creator. Once documents are made or received by the creator, the record-making activities outlined below must be carried out as a precursor to setting aside the records in the recordkeeping system and maintaining their authenticity.

Monitoring the Performance of the Record-making System (A2.1)

This activity involves assessing the efficacy of the performance of the record-making system as a whole by analyzing performance reports on the operation of each of the record-making system's records capture, identification, declaration, execution and transfer sub-systems and, in response, issuing (1) activity directives for record-making activities for each sub-system—that is, authoritative procedural orders/instruments intended to facilitate effective, co-ordinated and responsive record-making system activities—and (2) reports on the performance of the record-making system to the overarching performance monitoring function (Maintain Framework, A1.4) for use in continued refinement and maintenance of the chain of preservation framework.

Managing the Making and Receipt of Records (A2.2)

This activity provides overall control and co-ordination of the creator's activities related to capturing and identifying make internal documents and documents received from external juridical or physical persons and subsequently declaring and executing them as records. Specifically, the component activities of managing the making and receipt of records are:

1. *to make documents* (A2.2.1), which involves compiling or composing digital information in a syntactic manner in accordance with the specifications of the creator's documentary forms, integrated business and documentary procedures and record-making access privileges. This activity results in made documents, which are discrete aggregations of digital information that have been compiled in a syntactic manner, but which have not yet been recorded or "captured" (i.e., affixed to a digital medium with fixed form and stable content). In practical terms, this activity refers to document composition processes, such as writing correspondence, filling out forms or the process of compiling a document using information extracted from other sources, such as databases, documents and/or records;
2. *to capture documents made or received by the creator* (A2.2.2), which involves recording and saving (i.e., *affixing* to a digital medium in a *stable* syntactic manner) particular instantiations of incoming documents from external juridical or physical persons and internal documents made by the creator in the record-making system in accordance with the specifications of the creator's integrated business and documentary procedures and record-making access privileges, and attaching to each captured document the following metadata:
   a. For captured made documents:
      - chronological date (and possibly time) of compilation and capture;
      - documentary form—that is, whether the document is a report, a letter, a contract, etc.; and
      - digital presentation—that is, file format, wrapper, encoding, etc.

b. For captured received documents:
   - chronological date (and possibly time) of transmission from the originator;
   - chronological date (and possibly time) of receipt and capture;
   - documentary form—that is, whether the document is a report, a letter, a contract, etc.; and
   - digital presentation—that is, file format, wrapper, encoding, etc.

   Although, as modeled here, this is the first activity that generates "documents" in the InterPARES sense of the term—that is, as "indivisible units of information constituted by a message affixed to a medium (recorded) in a stable syntactic manner [with] fixed form and stable content"—in practice, the distinction between the previous "making" and this "capturing" activity will, in many situations, be essentially conceptual and transparent to users.

3. *to identify captured documents* (A2.2.3), which involves attaching to each captured document identity metadata that convey the action in which the document participates and its immediate context, and that are vital to enabling the presumption of authenticity of digital records,[20] including:
   - names of the persons involved in the creation of the document, including:
     - *author(s)*—that is, the physical or juridical person(s) responsible for issuing the document;
     - *writer(s)*—that is, the physical person(s) or position(s) responsible for articulating the content of the document;
     - *addressee(s)*—that is, the physical or juridical person(s) for whom the document is intended;
     - *originator(s)* (if different from the author or writer)—that is, the physical person(s), position(s) or office(s) responsible for the electronic account or technical environment where the document is generated and/or from which the document is transmitted;[21] and
     - *receiver(s) or recipient(s)*—that is, the physical or juridical person(s) to whom the document may be copied or blind copied for information purposes;
   - name of the action or matter—that is, the subject line(s) and/or the title at the top of the document;
   - indication of the presence of a digital signature;
   - indication of other forms of authentication, including:
     - *corroboration*—that is, an explicit mention of the means used to validate the document;
     - *attestation*—that is, the validation of the document by those who took part in the issuing of it, and by witnesses to the action or to the 'signing' of the document;
     - *subscription*—that is, the name of the author or writer appearing at the bottom of the document; and

[20] See Authenticity Task Force (2002), "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records," in Duranti, *Long-term Preservation* , op. cit., 204-219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf.

[21] Identification of the originator is only important in cases where the person, position or office responsible for physically creating and/or transmitting the document is neither the author nor the writer, and when the presence of the originator's name appearing on, or in association with, the document calls into question the actual author and/or writer of the document. This is most commonly associated with e-mails in instances where the name of the originator appears in the header of an e-mail and/or its attachments that were in fact authored and/or written by another person, but physically manifested and/or transmitted on behalf of that person by the originator.

  - *qualification of signature*—that is, the mention of the title, capacity and/or address of the person or persons signing the document;
- indication of any attachments—that is, mention of autonomous digital objects linked inextricably to the document.

4. *to declare captured and identified documents as records* (A2.2.4), which involves intellectually setting aside captured and identified documents as records by assigning classification codes from the classification scheme to the documents and adding these codes to the identifying metadata, and by assigning to the documents registration numbers based on the registration scheme and adding these numbers to the identifying metadata. The thinking here is that records classification, in particular, is a critical act in creating a record. It establishes that a document has been "set aside" for incorporation in the records of the creator and places it in relation to the business process of which it is a part, to the action that generated it and to other records. In some classification systems, it may also relate the record to the place in the administrative structure where it was created. In particular, a classification scheme or plan lays out what the aggregations of records are—that is, the series and classes of records—so that all of the records bearing on particular processes and matters can be identified and circumscribed. Identity metadata captured for this activity could include:
   New metadata:
   - classification code; and
   - registration number.

   Inherited metadata (i.e., identity metadata inherited from the current classification level and from all higher levels in the classification system, as applicable):
   - name of creator;
   - indication of copyright or other intellectual rights;
   - name of handling office (if not evident in the classification code);
   - name of office of primary responsibility (if not evident in the classification code and records retention schedule);
   - access restriction code (if not evident in the classification code);
   - access privileges code (if not evident in the classification code);
   - vital record code (if not evident in the classification code); and
   - planned disposition (if not evident in the classification code).
5. *to execute records* (A2.2.5), which involves attaching to each record metadata that convey information related to, and actions taken during the course of, the formal execution phase of the administrative procedure in which the record participates. This activity may also involve transmitting documents to external physical or juridical persons and making record copies of the sent documents. Metadata captured for this activity could include:
   - priority of transmission;
   - transmission date, time and/or place;
   - actions taken;
   - dates and times of further action or transmission; and
   - information on any attachments—that is, mention of autonomous items that were linked inextricably to the document prior to its transmission for the document to accomplish its purpose.

Managing the Setting Aside of Completed Records (A2.3)

This activity provides overall control and co-ordination of the transfer of completed (executed) records to the recordkeeping system. Specifically, the component activities of managing the setting aside of completed records are:

1. *to monitor the performance of the record-making transfer system* (A2.3.1), which involves assessing the efficacy of the performance of the record-making transfer system by analyzing reports on the operation of record-making transfer activities and, in response, issuing (1) activity directives for transfer activities and (2) reports on the performance of the record-making transfer system for use in continued maintenance of the record-making system;
2. *to prepare completed records for transfer to the recordkeeping system* (A2.3.2), which involves attaching to completed records integrity and related metadata that convey information related to, and actions taken during the course of, managing the records for records management purposes prior to setting them aside in the recordkeeping system; compiling information about the records that is needed to meet all transfer information requirements; and ensuring that the records are in the proper format for transfer to the recordkeeping system as prescribed by recordkeeping system rules and procedures and technological requirements. Metadata captured for this activity could include:
   - archival or filing date—that is, the date on which a record is officially incorporated into the creator's records;
   - draft or version number;
   - expression of archival bond (e.g., via classification code, file identifier, record item identifier, dossier identifier, etc.);
   - name of the creator—that is, the name of the physical or juridical person in whose archival fonds the record exists;
   - indication of copyright or other intellectual rights (if applicable);[22]
   - indication, as applicable, of the existence and location of duplicate records, whether inside or outside the record-making or recordkeeping systems and, in instances where duplicate records exist, which is the *authoritative copy*—that is, the instantiation of a record that is considered by the creator to be its official record and is usually subject to procedural controls that are not required for other instantiations;[23]
   - name of the handling office (if not evident in the classification code)—that is, the person or office using the record to carry out business;
   - name of the office of primary responsibility (if not evident in the classification code or the records retention schedule)—that is, the office given the formal competence for maintaining the authoritative version or copy of records belonging to a given class within a classification scheme;[24]
   - indication of any technical changes to the records—for example, change of encoding, wrapper or format, upgrading from one version to another of an application, or conversion of several linked digital components to one component only—by

[22] If a record comprises material copyrighted by different authors, indication of copyright clearance (or lack thereof) with related dates is necessary.

[23] InterPARES 2 Terminology Database, available at http://www.interpares.org/ip2/ip2_terminology_db.cfm. In cases where a record is certified by the author or creator as an "approved reproduction" of a work (for example, a digital work of art), indication of the existence of such certification is required.

[24] This may be the same as the handling person/office.

embedding directly in the record digital components that were previously only linked to the record, such as audio, video, graphic or text elements like fonts;
- indication of any annotations[25] or new attachments (e.g., records profiles);
- access restriction code (if applicable and if not evident in the classification code)—that is, indication of the person, position or office authorized to read the record;
- access privileges code (if applicable and if not evident in the classification code)—that is, indication of the person, position or office authorized to annotate the record, delete it, or remove it from the system;
- vital record code (if applicable and if not evident in the classification code)—that is, indication of the degree of importance of the record to continue the activity for which it was created or the business of the person/office that created it;[26] and
- planned disposition (if not evident in the classification code)—for example, removal from the live system to storage outside the system, transfer to the care of a trusted custodian, or scheduled deletion.

3. *to transfer completed records to the recordkeeping system* (A2.3.3)*,* which involves sending or transmitting completed records prepared for transfer to the office responsible for the recordkeeping function with the accompanying documentation necessary for recordkeeping. Metadata captured for this activity could include:
   - indication of the record(s) transferred;
   - name of the person effecting the transfer;
   - name of the entity to whom the records are transferred (if different than the office of primary responsibility); and
   - date/time of the transfer.

Managing Records in a Recordkeeping System (A3)

This third main activity involves overseeing and coordinating all the activities associated with maintaining records in the recordkeeping system to ensure their continuing authenticity, facilitating access to them, carrying out their disposition and monitoring the overall performance of the recordkeeping system.

Monitoring the Performance of the Recordkeeping System (A3.1)

This oversight activity involves assessing the efficacy of the recordkeeping system as a whole by analyzing reports on the operation of each of the recordkeeping system's records information, indexing, storage, retrieval, access and disposition sub-systems (and, as necessary, by examining records) and, in response, issuing (1) activity directives for each of the sub-systems and (2) reports on the performance of the recordkeeping system to the overarching performance monitoring function (Maintain Framework, A1.4) for use in continued refinement and maintenance of the chain of preservation framework.

Managing the Maintenance of Kept Records (A3.2)

This activity involves managing information about maintenance activities carried out on records in the recordkeeping system, managing the indexing of the records and the development

[25] Annotations are additions made to a record after it has been completed or executed. Therefore, annotations are not considered elements of the record's documentary form.

[26] The vital record code only pertains to specific communities of practices, such as legal and medical offices, who must identify the records that are vital to the continuance of their business in case of disaster and who would therefore exercise special protection measures on those records.

of indexing instruments and managing the storage of the records or, more accurately, their digital components in the recordkeeping storage system.

Managing Information About Kept Records (A3.2.1)

The main infrastructural component supporting this activity is a recordkeeping information system. It is important to emphasize that, as envisioned in this model, an information system (whether in the recordkeeping or permanent preservation system) is a set of rules governing the management and maintenance of information about the operation of the system and about the records in the system—including their digital components and the maintenance actions applied to them—and the tools and mechanisms used to implement these rules. Thus, it is much more than a software application, such as an electronic document and records management system (EDRMS)—which would, instead, simply constitute one of the tools within the system.

In general terms, this activity provides overall control and co-ordination of contextual and related information about records transferred to, and maintained in, the recordkeeping system, information about their access and use and information about ongoing records maintenance activities for use in records appraisal activities by the preserver and in records indexing, storage, access and disposition activities by the creator.

In terms of the underlying concern for maintaining the authenticity of the creator's records, the aim of this activity is to make sure that the nature of any actions undertaken upon the records is documented, whether through additions of integrity metadata or by compilations of reports, to provide a kind of audit trail on what has happened to the records since their creation. Such information is necessary when assessing the ongoing trustworthiness (reliability and authenticity) of the records in the system. In some cases, it may also be necessary to test the accuracy of records, such as when changes are made to a database that may affect the accuracy of data used to generate records.

The primary outputs of this activity include:

1. *recordkeeping information system performance information*—that is, continuously logged and updated documentation concerning the ability of the recordkeeping information sub-system to fulfil is purpose and achieve its performance objectives;
2. *information about kept records for appraisal*—that is, documentation compiled about the identity, integrity, format, form, context or other characteristics of records in the recordkeeping system for the purpose of appraising records and making appraisal decisions;
3. *information about kept records for creation*—that is, documentation compiled about records in the recordkeeping system for the purpose of helping inform and direct records creation activities;
4. *information about context*—that is, documentation compiled about the juridical-administrative, provenancial, procedural, documentary and/or technological contexts of kept records that is not available from the records themselves, for the purpose of facilitating appraisal;
5. *information about kept records in storage*—that is, documentation compiled about kept records in the recordkeeping storage system for the purpose of processing retrieval requests for records and/or information about records;
6. *information about digital components of kept records*—that is, technical information compiled about digital components of records in the recordkeeping storage system for the purpose of facilitating discovery of, and/or processing access requests for, records and/or information about records;

7. *information for indexing*—that is, documentation compiled about kept records for the purpose of establishing access points and creating indexing instruments to facilitate record discovery and retrieval; and
8. *information for storage of kept records*—that is, documentation compiled about kept records and their elements and digital components for the purpose of facilitating their storage and continued maintenance.

Managing the Indexing of Kept Records (A3.2.2)

This activity provides overall control and co-ordination of records indexing activities. The component activities of managing the indexing of kept records are:

1. *to monitor the performance of the recordkeeping indexing system* (A3.2.2.1), which involves assessing the efficacy of the performance of the recordkeeping indexing system by analyzing reports on the operation of recordkeeping indexing activities and, in response, issuing (1) activity directives for indexing activities and (2) reports on the performance of the indexing system for use in continued maintenance of the recordkeeping system;
2. *to index kept records* (A3.2.2.2), which involves establishing and recording access points for kept records within the context of a controlled recordkeeping vocabulary applied according to recordkeeping indexing system rules, procedures and strategies; and
3. *to develop indexing instruments* (A3.2.2.3), which involves preparing tools that facilitate discovery and retrieval of the records in the recordkeeping system, such as guides, inventories and indexes.

Managing the Storage of Kept Records (A3.2.3)

This activity provides overall control and co-ordination of the recordkeeping storage system and the records stored in the system. The component activities of managing the storage of kept records are:

1. *to monitor the performance of the recordkeeping storage system* (A3.2.3.1), which involves assessing the efficacy of the performance of the recordkeeping storage system by analyzing reports on the operation of recordkeeping storage activities and, in response, issuing (1) activity directives for storage activities and (2) reports on the performance of the storage system for use in continued maintenance of the recordkeeping system;
2. *to place the kept records in storage* (A3.2.3.2), which involves placing the digital components of kept records and their metadata into storage in accordance with the procedures for maintaining authentic records and the actions prescribed by the recordkeeping storage system strategies, rules and procedures and activity directives. Integrity metadata captured for this activity could include:
   - indication of the original state (e.g., file format) of the record(s) prior to storage;
   - indication of any modification(s) made to the record(s) in preparation for storage;
   - indication of the state of the record(s) after the modification(s) (e.g., impact on form, format, authenticity, etc.);
   - reason/authorization for the modification(s) (e.g., through reference to the relevant section of the recordkeeping storage system strategy);
   - date/time of any modification(s);
   - name of the person responsible for the modification(s);
   - name of person responsible for placing the record(s) in storage;

- date/time the record(s) was/were placed in storage; and
- location of the record(s) in storage.

3. *to maintain records in the recordkeeping storage system* (A3.2.3.3), which involves the following activities:
   a. *to monitor the kept records in storage* (A3.2.3.3.1), which involves (1) keeping track of the condition and maintenance requirements of kept records—more specifically, their digital components and metadata—and the media on which they are stored in the recordkeeping storage system to identify storage that needs backing-up, digital components and/or metadata that need correcting or updating and media that need refreshing and (2) issuing on maintenance activity reports that provide continuously updated documentation indicating the location of digital components of kept records in storage, the presence, nature and locations of recordkeeping system backups, the occurrence of storage problems, the actions taken to correct storage problems, the actions taken to update records and refresh storage media, the results of such actions and assessment of the impact, if any, of the maintenance activities on the authenticity of the records;
   b. *to back-up the recordkeeping storage system* (A3.2.3.3.2), which involves routinely creating copies of the digital content in the recordkeeping storage system for the purpose of recovery in the event of a disaster resulting in system failure or corruption, and recording information about the back-up activities. It is important to distinguish here between comprehensive *system backups* and localized *content backups*. System backups contain a copy of *all* the digital objects in the system, including the operating system, the software applications and all digital objects (i.e., digital components of the records and their metadata) in the system. A system backup provides the maximum level of recovery potential in the event of a disaster or system corruption. Content backups contain a copy of selected aggregations of the digital objects in the system and, therefore, only offer limited recovery potential, especially in cases where system applications become corrupted. Integrity metadata about these back-up activities, captured in activity reports, could include:[27]
      - indication of the reason/authorization for the backup (e.g., through reference to the relevant section of the recordkeeping storage system strategy);
      - indication of the type of backup (e.g., incremental, differential, full) ;
      - indication of the extent or content of the backup (e.g., full system, selected groups of records, etc.);
      - name of the person creating the backup;
      - date/time of the backup;
      - indication of the software application (including version number) used to create the backup; and
      - location of the backup; and
      - backup identification number.
   c. *to correct problems with kept records in storage* (A3.2.3.3.3), which involves taking the actions prescribed by the relevant recordkeeping storage system strategies, rules and procedures and activity directives, in accordance with the procedures for maintaining

[27] As per InterPARES 1 Benchmark Requirement A.3 - Protective Procedures: Loss and Corruption of Records (Authenticity Task Force, “Appendix 2,” op. cit., 211).

authentic copies of records, to identify and eliminate problems in storage to ensure that the records remain accessible, legible and intelligible over time; and recording information about the correction activities and the corrected digital components. Integrity metadata about this activity, captured either as metadata attached to the records or in activity reports, could include:[28]

- indication of the original state (e.g., file format) of the record(s) prior to correction;
- indication of the correction processe(s) used;
- indication of the state of the record(s) after correction (e.g., impact on form, format, authenticity, etc.);
- indication of the reason/authorization for the correction(s) (e.g., through reference to the relevant section of the recordkeeping storage system strategy);
- name of the person responsible for the correction(s);
- date/time of the correction(s);
- location of the corrected record(s); and
- correction identification number.

d. *to update kept records in storage* (A3.2.3.3.3), which involves carrying out conversion actions on the digital components of kept records in storage in accordance with the procedures for maintaining authentic records and the actions prescribed by the relevant recordkeeping storage system strategies, rules and procedures and activity directives, to ensure the records remain accessible, legible and intelligible over time; and recording information about the updating activities and the updated digital components. Typical conversion activities might include migration, standard-ization or transformation to persistent form. Integrity metadata related to this activity would be similar to those noted above for correction activities;[29] and

e. *to refresh the media for kept records in storage* (A3.2.3.3.5), which involves copying or transferring the digital components of kept records in storage from one medium to another—or otherwise ensuring that the storage medium remains sound—in accordance with the procedures for maintaining authentic records and the relevant actions prescribed by the recordkeeping storage system strategies, rules and procedures and activity directives; and recording information about the refreshment activities and about any impact to the digital components on the refreshed media. Again, integrity metadata related to this activity would be similar to those noted above for correction activities.[30]

Managing Access to Kept Records (A3.3)

The recordkeeping access system is defined as a set of rules governing the specific methods and strategies for discovering, reconstituting and presenting and/or packaging retrieved records and/or information about records in the recordkeeping system and the tools and mechanisms used to implement these rules.

This activity entails facilitating the discovery of, and managing access requests for, kept records and/or information about kept records and monitoring the performance of the recordkeeping access system.

[28] Ibid.

[29] As per InterPARES 1 Benchmark Requirement A.4 - Protective Procedures: Media and Technology (Authenticity Task Force, "Appendix 2," op. cit.).

[30] Ibid.

Monitoring the Performance of the Recordkeeping Access System (A3.3.1)

This activity involves assessing the efficacy of the performance of the recordkeeping access system by analyzing reports on the operation of recordkeeping access activities and, in response, issuing (1) activity directives for access activities and (2) reports on the performance of the access system for use in continued maintenance of the recordkeeping system.

Facilitating the Discovery of Kept Records and/or Information (A3.3.2)

This activity provides authorized internal and external users with mediated records query, search and discovery support through the use of tools such as thesauri and records indexes.

Managing Requests for Kept Records and/or Information (A3.3.3)

This activity provides overall control and co-ordination of internal and external requests for access to records and/or information about kept records by processing access requests, retrieving the digital components for the requested records and/or information, verifying the retrieved components and information and providing access to the retrieved records and/or information. The component activities of managing requests for kept records and/or information are:

1. *to process requests for kept records and/or information* (A3.3.3.1), which involves the following activities:
   a. *to register recordkeeping access requests* (A3.3.3.1.1), which involves recording registration information about received mediated and unmediated requests for access to kept records and/or information about the records and issuing notifications of receipt to the persons requesting the records. Integrity metadata captured for this activity could include:
      - name of the person requesting the record(s) and/or information;
      - name of the person for whom the request is being made (if different than the above);
      - date/time of the request;
      - indication of the records and/or information requested;
      - access privileges of the requester;
      - name of the person registering the request;
      - request registration number;
      - indication of notification of receipt sent (including indication of any additional information needed to register the request, if any);
      - name of person to whom the notification of receipt was sent;
      - name of the person issuing the notification of receipt; and
      - date/time the notification of receipt was sent.
   b. *to retrieve information needed to process recordkeeping access requests* (A3.3.3.1.2), which involves gathering the information, from indexing instruments, controlled vocabularies, record profiles and other recordkeeping tools, needed to process the access requests and translate them into retrieval requests;
   c. *to generate recordkeeping retrieval requests* (A3.3.3.1.3), which involves translating access requests into requests to the recordkeeping storage and information systems for retrieval of the exact digital components and/or information required to fulfil the access requests; and
   d. *to generate recordkeeping requests specifications* (A3.3.3.1.4), which involves issuing instructions to the recordkeeping retrieval and access systems on how to fulfil requests for kept records and/or information about the records based on analyses of the requests and the processing information received from A3.3.3.1.3 in relation to

the relevant recordkeeping access system strategies, rules and procedures (including procedures for maintaining authentic records) and access privileges;

2. *to retrieve the requested kept records and/or information* (A3.3.3.2), which involves outputting copies of the digital components of the requested records and information about the records and their digital components—such as their identity, integrity, format, form, context, content, etc.—in response to retrieval requests; and
3. *to verify the retrieved kept records and/or information* (A3.3.3.3), which involves determining (1) whether all the components and information necessary to satisfy an access request have been received from the recordkeeping storage and information systems, (2) whether the retrieved components and information can be processed for output and (3) in cases where digital components are encountered that need updating or correcting, redirecting them (or information about the problems encountered) to the maintenance function of the recordkeeping storage system. Integrity metadata captured for this activity could include:
   - request registration number;
   - indication of the measures used to verify the retrieved digital components and/or information;
   - name of the person verifying the retrieved digital components and/or information;
   - indication of the determination of verification (i.e., verified, rejected);
   - reason(s) for rejection (as appropriate);
   - indication of required maintenance action(s) (as appropriate); and
   - date of verification/rejection.
4. *to provide access to the retrieved kept records and/or information* (A3.3.3.3), which involves providing users with access to copies of kept records and/or information about the records. For certain requests, such as those under freedom of information and privacy laws, it may be necessary to document and keep information about the records and/or information issued in response to requests, to whom the records and/or information were issued and when. In cases where redacted records are issued, a copy of the redacted[31] record should be created and its existence properly documented with the appropriate metadata. Unsuccessful attempts to provide access to the records and/or information trigger the creation of a "notification of rejection of recordkeeping access request" that is issued to the requestor, a record copy of which is also retained by the creator. The component activities of providing access to retrieved kept records and/or information are:
   a. *to reconstitute kept records and/or information* (A3.3.3.3.4.1), meaning to link or assemble all the verified digital components of the requested records and/or information for the purpose of reproducing and manifesting (presenting) to the user copies of the records and/or information in authentic form and, if necessary, to redact information to meet privacy and/or copyright requirements. Integrity metadata captured for this activity could include:
      - indication of any problems encountered in reconstituting the records and/or information in authentic form;
      - indication of required maintenance action(s);
      - indication of any redaction for privacy or copyright reasons;
      - indication of the reason/authorization for the redaction;

[31] Richard Pearce-Moses, in his *A Glossary of Archival and Records Terminology*, defines redaction as "the process of concealing sensitive information in a document before being released to someone not authorized to see that information." See http://www.archivists.org/glossary/index.asp.

- date of the redaction;
- name of the person responsible for handling/executing the redaction; and
- registration number of the record copy of the redacted record issued to the user.

b. *to manifest kept records and/or information* (A3.3.3.3.4.2), meaning to present to the user copies of the reconstituted requested records and/or information about the records with the appropriate extrinsic form and, in the case of records aggregates, with information about their relationships to one another (archival bond). This activity also involves providing users with a Certificate of Authenticity, if requested. Regarding metadata, this activity results in the production or compilation of two sets: one set of integrity metadata for the creator to document the activity and one set of identity and integrity metadata for the user. Metadata captured for this activity could include:

Records Creator (integrity metadata)

For requests that are fulfilled (in part or in whole)

- indication of the record(s) and/or information presented;
- indication of any redaction for privacy or copyright reasons (as appropriate);
- indication of a Certificate of Authenticity, if issued;
- indication of the means by which the records were authenticated[32]
- name of the person to whom the record(s) and/or information were presented;
- date when the record(s) and/or information were presented;
- name of the person responsible for handling/effecting the access request;
- indication of the state or condition of the record(s) and/or information at time the request was fulfilled (including, especially, an indication of instances where a copy of a presented record is known not to fully and faithfully reproduce the elements expressing its identity and integrity);[33] and
- indication of any problems encountered in manifesting the records and/or information in authentic form.

For requests that cannot be fulfilled (in part or in whole)

- indication of why the request cannot be fulfilled;
- name of the person responsible for determining that the request cannot be fulfilled;
- indication that a notification of rejection was sent;
- name of the person to whom the rejection notification was sent;
- name of the person responsible for issuing the rejection notification; and
- date/time the rejection notification was sent.

Records User

Identity metadata[34]

- name(s) of the person(s) concurring in formation of the record(s);
- name(s) of action or matter;
- date(s) of creation and transmission of the record(s);
- expression of archival bond; and

[32] As per InterPARES 1 Benchmark Requirement A.6 - Authentication of Records (Authenticity Task Force, "Appendix 2," op. cit., 212).
[33] In reference to InterPARES 1 Baseline Requirement B.2.d. (Ibid., 213.).
[34] As per InterPARES 1 Benchmark Requirement A.1 - Expression of Record Attributes and Linkage to Record (Ibid., 210.).

- indication of any attachments.

Integrity metadata (as necessary)

- indication of access privileges used to control creation and maintenance of the presented record(s);[35]
- indication of protective procedures used to prevent corruption of the presented record(s);[36]
- indication of protective procedures used to guarantee the continuing identity and integrity of the presented records against media deterioration and across technological change;[37]
- indication of the means by which the presented record(s) was/were authenticated;
- indication of instances where a copy of a presented record is known not to fully and faithfully reproduce the elements expressing its identity and integrity;
- indication of any redaction for privacy or copyright reasons;
- indication of the reason/authorization for the redaction;
- date of the redaction;
- name of the person responsible for handling/executing the redaction;
- date when the requested record(s) and/or information were presented; and
- name of the person responsible for handling/executing the access request.

c. *to package kept records and/or information for output* (A3.3.3.4.3), meaning to combine the digital components of the requested records and/or information with instructions on how to reconstitute and manifest the records or information with the appropriate extrinsic form. Regarding metadata, this activity results in the production of the same two sets of metadata outlined above, with the exception of the indication of a Certificate of Authenticity.

Managing the Disposition of Kept Records (A3.4)

This final recordkeeping system activity provides overall control and co-ordination of records disposition activities, including monitoring the performance of the disposition system, processing disposition information and, in accordance with disposition activity directives and disposition rules, procedures and strategies, destroying kept records and/or preparing and transferring kept records to the designated preserver. The component activities of managing the disposition of kept records are:

1. *to monitor the performance of the disposition system* (A3.4.1), which involves assessing the efficacy of the performance of the recordkeeping disposition system by analyzing reports on the operation of disposition activities and feedback/information received via records transfer notifications[38] and, in response, issuing (1) activity directives for disposition activities and (2) reports on the performance of the disposition system for use

[35] As per InterPARES 1 Benchmark Requirement A.2 - Access Privileges (Ibid., 211.).

[36] As per InterPARES 1 Benchmark Requirement A.4 - Protective Procedures: Media and Technology (Ibid.).

[37] Ibid.

[38] These transfer notifications include: *Notifications of Receipt of Transfer*, which are formal instruments sent to the creator by the preserver acknowledging that the preserver has received the transfers and, if needed, requesting that the creator address any problems encountered in registering the transfers; and *Notifications of Rejection of Transfer*, which are formal instruments sent to the creator by the preserver indicating that transfers of records do not satisfy requirements for being accessioned or preserved, because the transfers are unauthorized, do not contain the proper records, contain records that cannot be authenticated or whose preservation is not feasible. For further details about these instruments, see discussion of activity 4.3. Acquiring Selected Records, below.

in continued maintenance of the recordkeeping system. A specific type of activity directive related to this activity would be an order to rectify a rejected transfer in response to receipt of a "notification of rejection of transfer" from the designated preserver. Such a directive would instruct disposition activity management staff to remedy problems that resulted in rejection of an attempted records transfers and, as appropriate, reinitiate the transfer;

2. *to identify kept records for disposition* (A3.4.2), which involves identify records and information about records in the recordkeeping system that are earmarked either for destruction or for transfer to the designated preserver, as determined by the creator's retention schedule;
3. *to destroy kept records* (A3.4.3), which involves obliterating the kept records, and information related to the records (e.g., record profiles, index references, etc.), identified for destruction and providing documentation about the records destroyed. Metadata captured for this activity could include:[39]
   - indication of the records and related information (e.g., records profiles, index references, etc.) destroyed;
   - indication of the reason/authorization for the destruction (e.g., reference to the relevant retention schedule, including the version number of the retention schedule, as applicable);
   - name of the person responsible for handling/executing the destruction; and
   - date/time of the destruction.
4. *to prepare kept records for transfer to the designated preserver* (A3.4.4), which involves attaching to the kept records identified for transfer integrity and related metadata about actions taken during the course of preparing the records for transfer to the designated preserver in accordance with the terms and conditions of transfer, and compiling information about the records that is needed to meet all information requirements of the designated preserver, as outlined in the terms and conditions of transfer. Integrity metadata captured for this activity could include:[40]
   - indication of any technical changes applied to the records in preparation for the transfer (e.g., conversion to a new format), including the results/consequences of the actions (especially with regard to authenticity);
   - indication of the reason/authorization for the actions (e.g., reference to the relevant terms and conditions of transfer);
   - name of the person responsible for handling/executing the transfer preparation actions; and
   - date/time when the actions were carried out.
5. *to transfer kept records to the designated preserver* (A3.4.5), which involves relocating kept records selected for long-term preservation to a designated records preserver (or, as applicable, the office of the creator responsible for the permanent preservation function), along with the accompanying transfer documentation necessary for permanent preservation. Such documentation includes administrative information about the transfer needed by the preserver to register the transfer, confirm the authorization for the transfer and verify its contents—such as an indication of the entity transferring records, the

[39] As per InterPARES 1 Benchmark Requirement A.8 - Removal and Transfer of Relevant Documentation (Authenticity Task Force, "Appendix 2," op. cit., 212.).
[40] Ibid.

contents of the transfer, the terms and conditions governing the transfer, etc. It also includes more specific information about the records in the transfer that the preserver will need for the purposes of: (1) establishing the identity and demonstrating the integrity of the records being transferred, (2) identifying their logical format, constituent digital components, documentary form and other preservation-related characteristics, (3) properly ordering the records with respect to their relationships with each other (i.e., their archival bond) and (4) associating the records with their relevant contexts (juridical-procedural, provenancial, procedural, documentary, technical). Integrity metadata captured for this activity could include:[41]

- indication of the records transferred;
- indication of the reason/authorization for the transfer (e.g., reference to the relevant terms and conditions of transfer);
- creator's transfer registration number;
- name of the person responsible for handling/executing the transfer;
- name of the entity to whom the records were transferred; and
- date/time of the transfer.

Managing Records in a Permanent Preservation System (A4)

This third main activity involves overseeing and coordinating all the activities associated with preserving records in the permanent preservation system to ensure their continuing authenticity while in the custody of the designated preserver. The key activities of the permanent preservation function are to appraise and select, acquire, preserve and output records selected for long-term preservation, and to monitor the performance of the permanent preservation system. These activities are viewed from the perspective of the entity responsible for long-term preservation of authentic copies of the creator's digital records; that is, the designated records preserver, who will carry out the preservation activities. By contrast, the activities of record-making and recordkeeping, discussed above, were viewed from the perspective of the preserver, but as carried out by the records creator. It is true that records creators may also take responsibility for long-term preservation of digital records. In such a case, the records creator would still have to adopt the perspective (and responsibilities) of the designated records preserver to undertake the following activities.

Monitoring the Performance of the Permanent Preservation System (A4.1)

Similar the performance monitoring function for the recordkeeping system, this activity involves assessing the efficacy of the permanent preservation system as a whole by analyzing reports on the operation of each of the permanent preservation system's records information, selection, acquisition, description, storage, retrieval and access sub-systems (and, as necessary, by examining records) and, in response, issuing (1) activity directives for each of the sub-systems and (2) reports on the performance of the permanent preservation system to the overarching performance monitoring function (Maintain Framework, A1.4) for use in continued refinement and maintenance of the chain of preservation framework.

Appraising Records for Permanent Preservation (A4.2)

Although appraising records is ultimately the responsibility or preservers, in cases where, as recommended in this model, retention scheduling is employed, decisions on the disposition of

[41] Ibid.

records will regularly be made by a records creator as part of the design, implementation and management of its recordkeeping system. In some cases, appraisals may be made when it is determined that records in a longstanding system need to reach a disposition. In either case, this activity entails making appraisal decisions by compiling information about kept records and their context, assessing their value, determining the feasibility of their preservation and monitoring the performance of the preservation selection system. It also involves monitoring appraised records and appraisal decisions, in relation to updated information about the creator and the preserver, to identify evolving conditions that might make it necessary for the preserver to adjust or redo an appraisal.

Monitoring the Performance of the Preservation Selection System (A4.2.1)

This activity involves assessing the efficacy of the performance of the preservation selection system by analyzing reports on the operation of preservation selection activities and, in response, issuing (1) directives for selection activities and (2) reports on the performance of the selection system for use in continued maintenance of the permanent preservation system.

Analyzing Kept Records for Preservation (A4.2.2)

This activity involves assessing information concerning the kept records being appraised, including information about their contexts, value and preservation feasibility. In some cases, this activity could be initiated through instructions issued by the appraisal monitoring function (discussed below) to revise or update previous appraisal decisions. In particular, the component activities of analyzing kept records for preservation are:

1. *to analyze information about appraised records* (A4.2.2.1), which involves collecting, organising, recording and assessing relevant information from the kept records being appraised and about their juridical-administrative, provenancial, procedural, documentary and technological contexts;
2. *to assess the value of appraised records* (A4.2.2.2), which involves the following activities:
   a. *to assess the continuing value of the appraised records* (A4.2.2.2.1), which means to determine the capacity of the records being appraised to serve the continuing interests of their creator and society;
   b. *to assess the authenticity of the appraised records* (A4.2.2.2.2), which means to determine the grounds for presuming the records to be authentic. This activity, in turn, entails: (1) compiling evidence to support a presumption of authenticity (A4.2.2.2.2.1), which involves collecting, organizing and recording evidence of the identity and integrity of the records being appraised and about the procedural controls applied to them during their creation and maintenance by the creator, (2) measuring the evidence compiled about the identity, integrity and procedural controls of the records being appraised against the requirements for authentic records[42] (A4.2.2.2.2.2) and (3) in cases where the examined evidence is too weak to support a presumption of authenticity, using other verification methods[43] to determine the

[42] The model assumes the requirements as listed in the InterPARES 1 Project's "Requirements for Assessing and Maintaining the Authenticity of Electronic Records" (Authenticity Task Force, "Appendix 2," op. cit., 204–219).

[43] Alternative "[m]ethods of verification include, but are not limited to, a comparison of the records in question with copies that have been preserved elsewhere or with back-up tapes; comparison of the records in question with entries in a register of incoming and outgoing records; textual analysis of the record's content; forensic analysis of aspects such as medium and script; a study of audit trails; and the testimony of a trusted third party" (Heather MacNeil et al., "Part One – Establishing and Maintaining Trust in

authenticity of the records being appraised in cases where there is insufficient evidence to meet the requirements for presuming the authenticity of the records; and

c. *to determine the value of the appraised records* (A4.2.2.2.3), which means to establish the value of the appraised records based on the results of the assessments of the records' authenticity and continuing value as well as information about their suitability and relevance in relation to the preserver's mission and its existing holdings;

3. *to determine the feasibility of preserving the appraised records* (A4.2.2.3), which involves the following activities:
   a. *to determine the record elements to be preserved* (A4.2.2.3.1), which means to identify, among the records being appraised, the necessary documentary components (e.g., record profile, attachments, annotations, etc.) and elements of form (e.g., author, date, subject line, etc.) that must be preserved to protect the authenticity of those records;
   b. *to identify the digital components to be preserved* (A4.2.2.3.2), which means to identify the digital components that manifest the record elements that need to be preserved to protect the authenticity of the records selected for permanent preservation; and
   c. *to reconcile the identified preservation requirements with the preserver's preservation capabilities* (A4.2.2.3.3), which means to determine whether the digital components manifesting the record elements that need to be preserved to protect the authenticity of the records selected for permanent preservation can in fact be preserved given the preserver's current and anticipated preservation capabilities.

The final outputs of these three activities include: (1) documentation about the digital components to be preserved—specifically, information about the way in which the record elements to be preserved are manifested in the electronic environment, construed for the purposes of instructing preservation activities and (2) preservation feasibility reports—that is, documented assessments of whether the record elements and digital components of the records proposed for preservation can be preserved given the preserver's current and anticipated preservation capabilities.

Making Appraisal Decisions (A4.2.3)

This activity involves determining and documenting the retention and disposition of the selected records based on valuation and feasibility information and agreeing on and documenting—through consultation with the creator—the terms and conditions of transfer of the selected records to the preserver. InterPARES 2 defines *terms and conditions of transfer* as "formal instruments that identify in archival and technological terms digital records to be transferred, together with relevant documentation [for their long-term preservation], and that identifies the medium and format of transfers, when the transfers will occur, and the parties to the transfers."[44]

Monitoring Appraisal Decisions (A4.2.4)

Because there may be changes in the way records are generated or organized, in the technology the creator uses to create them, or in the preserver's preservation capabilities, part of

Electronic Records: Authenticity Task Force Report," in Duranti, *Long-term Preservation*, op. cit., 50. Online reprint available at http://www.interpares.org/book/interpares_book_d_part1.pdf).

[44] InterPARES 2 Project Terminology Database, op. cit.

appraising digital records involves monitoring records that have already been appraised to identify any necessary changes to appraisal decisions over time. As well, because the creator's organizational mandates and responsibilities may change over time, as might the way those responsibilities are carried out, such that data accumulated in formerly appraised systems may be put to new uses, it is possible that systems that did not initially contain records may be upgraded to do so, especially in organizations with hybrid paper and electronic recordkeeping systems. Likewise, it is likely that the preserver's preservation capabilities will change over time, as might its organizational mandates and responsibilities. Therefore, in addition to monitoring changes to the creator's appraised records, it is also necessary for the preserver to keep track of appraisal decisions in relation to subsequent developments within the creator's and/or preserver's operations that might make it necessary to adjust or redo an appraisal, such as substantial changes to: (1) the creator's organizational mandate and responsibilities, (2) the creator's record-making or recordkeeping activities or systems, (3) the preserver's records preservation activities or systems and/or (4) the preserver's organizational mandate and responsibilities.

Acquiring Selected Records (A4.3)

It is an assumption of the model that custody and control of digital records will move from the creator to the preserver. It is true that records creators often maintain digital records for a long time, and so face many of the problems of long-term preservation, particularly when records have to be removed from active recordkeeping systems. In this model, the activity of the preserver's acquiring selected records and all the activities of preservation that follow on from that have as their goal the continued accessibility and authenticity of those records that are selected for continuing preservation, that is, for which one does not see an end to their preservation. This movement of records from the creator's hands to the preserver's hands is a critical juncture, and involves taking great care to make sure nothing goes awry in the transfer process. Acquiring selected records entails processing records transfers, accessioning accepted transfers and monitoring the performance of the acquisition system.

Monitoring the Performance of the Preservation Acquisition System (A4.3.1)

This activity involves assessing the efficacy of the performance of the preservation acquisition system by analyzing reports on the operation of preservation acquisition activities and, in response, issuing (1) directives for acquisition activities and (2) reports on the performance of the acquisition system for use in continued maintenance of the permanent preservation system.

Processing Records Transfers (A4.3.2)

This model envisages a five-step processing process involving the following activities:

1. *to register the transfer* (A4.3.2.1), which involves recording information about the transfer to register the circumstances of its occurrence. Specifically, this activity involves capturing the following metadata:[45]
   - name of the person responsible for effecting the transfer;
   - transfer registration number assigned by the transferring agent;
   - date and time the transfer was received;
   - name of the person registering the transfer;

[45] As per InterPARES 1 Baseline Requirement B.1 - Controls over Records Transfer, Maintenance, and Reproduction (Authenticity Task Force, "Appendix 2," op. cit., 213).

- transfer registration number assigned by the person registering the transfer;
- indication of the reason/authorization for the transfer (e.g., reference to the relevant terms and conditions of transfer);
- indication of records and other transfer documentation received;
- name of person(s) to whom a notification of receipt of transfer was issued;
- name of the person who issued the notification; and
- date and time the notification was sent.

This information constitutes metadata about the records in the transfer and may be recorded in an electronic register as part of the preservation retrieval system. When a transfer has been registered as having been received, the recipient notifies the transferring agent that the transfer (by number) has been received, and records this act as well. At this stage, it should be noted, as will become obvious, that nothing is done to establish that the transfer is correct in every respect; instead, the process of registration involves simple acknowledgement of the receipt of the transfer, upon which the next step follows closely.

2. *to confirm the authorization for the transfer* (A4.3.2.2), which involves confirming the person transferring the records has the authority to transfer records selected for preservation, and, in cases of unauthorized persons effecting transfers, issuing notifications of rejection of transfer to the persons transferring the records.
   a. If the transfer is accepted as being authorized, the following information is recorded as metadata in the register of transfers:[46]
      - date/time the transfer was accepted as authorized;
      - name of the person confirming the authorization of transfer;
      - transfer authorization number (as assigned by the preserver); and
      - terms and conditions of transfer number.
   b. If the transfer is rejected, the following information is recorded in the register:
      - date/time the transfer was rejected as unauthorized;
      - name of the person rejecting the transfer;
      - name of person(s) to whom a notification of rejection of transfer was issued;
      - name of the person who issued the rejection notification;
      - date and time the rejection notification was sent;
      - indication of the reason for the rejection;
      - transfer authorization rejection number (as assigned by the preserver); and
      - terms and conditions of transfer number.

   Note that this procedure assumes that the preserver has a list/mechanism to identify persons authorized to effect transfers. In the electronic environment it is all too easy for transfers to come from unauthorized persons who may have a copy of records, so care is needed to ensure that transfers come from authorized sources.
3. *to verify the content of the transfer* (A4.3.2.3), which involves determining whether transfers of records selected for preservation have been successfully transmitted (i.e., were not corrupted during transmission) and include all records and aggregates of records specified in the terms and conditions of the transfer, and, in corrupted or unverified cases, issuing notifications of rejection of transfer to the persons transferring the records.
   a. If the content of the transfer is accepted as being correct, the following information is recorded as metadata in the register of transfers:[47]

[46] Ibid.

- date/time the transfer was accepted as verified;
- indication of the measures used to verify the transfer;
- name of the person verifying the transfer;
- transfer content verification number (assigned by the preserver); and
- terms and conditions of transfer number.

b. If the transfer is rejected, the following information is recorded in the register:
- date/time the transfer was rejected as containing incorrect or corrupted content;
- name of the person rejecting the transfer;
- name of the person(s) to whom a notification of rejection of transfer was issued;
- name of the person who issued the rejection notification;
- date and time the rejection notification was sent;
- indication of the measures used to assess the content of the transfer;
- indication of the reason(s) for the rejection;
- transfer content rejection number (as assigned by the preserver); and
- terms and conditions of transfer number.

4. *to confirm the authenticity of the records in the transfer* (A4.3.2.4), which involves determine whether the assessment of the authenticity of the creator's records being transferred, which was conducted as part of the appraisal process, is still valid by verifying that the attributes relating to the records' identity and integrity have been carried forward with them along with any relevant documentation.

a. If the authenticity of the records confirmed, the following information is recorded as metadata in the register of transfers:[48]
- date/time the transfer was accepted as containing authentic records;
- indication of the measures used to confirm authenticity;
- name of the person confirming the authenticity;
- authenticity assessment report number (assigned by the preserver);
- transfer authenticity verification number (assigned by the preserver); and
- terms and conditions of transfer number.

b. If the authenticity of the records is not confirmed, the following information is recorded as metadata in the register of transfers:
- date/time the transfer was rejected as containing records that could not be authenticated;
- name of the person rejecting the transfer;
- name of the person(s) to whom a notification of rejection of transfer was issued;
- name of the person who issued the rejection notification;
- date and time the rejection notification was sent;
- indication of the measures used to assess the authenticity of the records in the transfer;
- indication of the reason(s) for the rejection;
- transfer authenticity rejection number (as assigned by the preserver);
- authenticity assessment report number (assigned by the preserver); and
- terms and conditions of transfer number.

[47] Ibid.
[48] Ibid.

5. *to confirm the feasibility of preserving the transfer* (A4.3.2.5), which involves verifying that the determination of the feasibility of preservation made during the process of appraisal is still valid and, in unconfirmed cases, results in issuance of notifications of rejection of transfer to the persons transferring the records. At this stage, before accessioning the records and formally accepting the records under the custody and control of the preserver, it must be confirmed that the preserver's current and expected future capabilities are sufficient to preserve the records over the long term. In particular, there may have been changes in the technology or assumptions made at the time of appraisal that no longer stand and invalidate the original feasibility assessment. This process also generates metadata in the register of transfers.
   a. If it proves feasible to preserve the records, the following information is recorded as metadata in the register of transfers:[49]
      - date/time the feasibility of preservation was confirmed;
      - name of the person confirming the feasibility;
      - feasibility report number (assigned by the preserver);
      - feasibility verification number (assigned by the preserver); and
      - terms and conditions of transfer number.
   b. If it proves not feasible to preserve the records, the following information is recorded as metadata in the register:
      - date/time the transfer was rejected as containing records that cannot be preserved;
      - name of the person rejecting the transfer;
      - name of the person(s) to whom a notification of rejection of transfer was issued;
      - name of the person who issued the rejection notification;
      - date and time the rejection notification was sent;
      - indication of the measures used to confirm the feasibility of preservation;
      - indication of the reason(s) for the rejection;
      - feasibility report number (assigned by the preserver);
      - feasibility rejection number (assigned by the preserver); and
      - terms and conditions of transfer number.

Accessioning Records (A4.3.3)

This activity, which only occurs after a transfer is registered, verified as coming from an authorized source and has had its content, authenticity and feasibility of preservation confirmed, involves formally documenting the acceptance of the transferred records into the custody of the preserver. The process of accessioning generates the following information to be recorded as metadata in the register of accessions:[50]

a. Identity metadata:
   - the records accessioned, including:
     - name of the juridical or physical person that created the records;
     - name of the juridical or physical person that transferred, donated or sold the records; and
     - quantity and characteristics of the records;
   - transfer registration number;

[49] Ibid.
[50] Ibid.

- accession registration number;
- accrual registration number (as appropriate);
- date the records are accessioned;
- indication of the digital rights that apply to the records accessioned, including:
  - name of the person(s) holding the rights;
  - terms and condition of the rights, including jurisdiction, duration, pertaining to which records, etc.; and
  - rights document number (e.g., deed of gift, contract, etc.);
- name of the person responsible for effecting the accession; and
- location of the accession.

b. Integrity metadata:
- original state of the records in the transfer when received;
- indication of the security and control procedures used for the transfer;
- indication of any modifications made to the records since their receipt
- indication of the post-modification state of the records (especially in relation to the impact of the modifications on the records' form, format, authenticity, etc.) (as appropriate);
- reason/authorization for the modifications (as appropriate);
- date of the modifications (as appropriate); and
- name of the person responsible for the modifications (as appropriate).

Preserving Accessioned Records (A4.4)

This function breaks down into three activities: (1) managing information about records acquired for permanent preservation, (2) managing the description of the records and (3) managing the storage of their digital components and related information.

Managing Information About Preserved Records (A4.4.1)

This activity is an important, complicated and vital facet of the process of preserving digital records. It involves compiling and updating information for all activities related to the preservation, description, storage, discovery, retrieval and output of records acquired for permanent preservation. Essentially, it is an exercise in management of information about preserved records and it is obvious that to achieve the ends of this facet of the process will require a robust information system to allow for the control, identification, retrieval, use and updating of the records as actions are taken upon them.

As with the recordkeeping information system, in terms of the underlying concern for maintaining the authenticity of the records in the custody of the preserver, the aim of this activity is to make sure that the nature of any actions undertaken upon the records is documented, whether through additions of integrity metadata or by compilations of reports, to provide a kind of audit trail on what has happened to the records since their creation. Such information is necessary when assessing the ongoing trustworthiness (reliability and authenticity) of the records in the system. In fact, careful, consistent and complete documentation of preservation activities is one of the fundamental requirements for ensuring the production of authentic copies of authentic records outlined in InterPARES 1 Baseline Requirement B.2 (Documentation of Reproduction Process and its Effects), which are, in turn, used to support Baseline Requirement B.3 (Archival Description). The critical nature of this documentation and the role that it plays in helping ensure the continued authenticity of the copies of the creator's records in the custody of the preserver is best expressed in the extended commentary to these two baseline requirements, which are reproduced below:

> *B.2 Documentation of Reproduction Process and its Effects*
> Documenting the reproduction process and its effects is an essential means of demonstrating that the reproduction process is transparent (i.e., free from pretence or deceit). Such transparency is necessary to the effective fulfilment of the preserver's role as a trusted custodian of the records. Documenting the reproduction process and its effects is also important for the users of records since the history of reproduction is an essential part of the history of the record itself. Documentation of the process and its effects provides users of the records with a critical tool for assessing and interpreting the record.[51]
> *B.3 Archival Description*
> Traditionally it has been a function of archival description to authenticate the records and perpetuate their administrative and documentary relationships. With electronic records, this function becomes critical. Once the records no longer exist except as authentic copies, the archival description is the primary source of information about the history of the record, that is, its various reproductions and the changes to the record that have resulted from them. While it is true that the documentation of each reproduction of the record copies[28] may be preserved, the archival description summarizes the history of all the reproductions, thereby obviating the need to preserve all the documentation for each and every reproduction. In this respect, the description constitutes a collective attestation of the authenticity of the records and their relationships in the context of the fonds to which the records belong. This is different from a certificate of authenticity, which attests to the authenticity of individual records. The importance of this collective attestation is that it authenticates and perpetuates the relationships between and among records within the same fonds.[52]

With this in mind, the focus can now turn to the component activities of managing the information about preserved records, which are:

1. *to monitor the performance of the preservation information system* (A4.4.1.1), which involves assessing the efficacy of the performance of the permanent preservation information system by analyzing reports on the operation of information management activities and, in response, issuing (1) activity directives for information activities and (2) reports on the performance of the information system for use in continued maintenance of the permanent preservation system;
2. *to compile information for preservation activities* (A4.4.1.2), which involves collecting, organizing and recording relevant appraisal, acquisition, accession and preservation information about acquired records to facilitate their preservation, arrangement, description, storage, discovery, retrieval and output. Effective work in this area will rationalize storage and retrieval of metadata and the information in documentation created during the processes of records creation and recordkeeping and while appraising, acquiring and accessioning records. Together all this information constitutes an important

[51] Authenticity Task Force, "Appendix 2," op. cit. 218.
[52] Ibid., 218–219. Footnote 28 in the original text reads: Although, technically, every reproduction of a record that follows its acquisition by the preserver is an authentic copy, it is the only record that exists and, therefore, should normally be referred to as "the record" rather than as "the copy."

source for the processes of arranging, describing, storing, locating, retrieving and outputting records. The key outputs of this activity include:

a. *preservation information system performance information*—that is, continuously logged and updated documentation concerning the ability of the permanent preservation information sub-system to fulfil is purpose and achieve its performance objectives;
b. *information about preserved records in storage*—that is, documentation compiled about preserved records in the permanent preservation storage system for the purpose of processing retrieval request for records and/or information about records;
c. *information about digital components of preserved records in storage*—that is, technical documentation concerning digital components of records in the preservation storage system that is needed to facilitate discovery of, and/or process access requests for, the records and/or information about the records;
d. *information about the preserver's existing holdings*—that is, documentation compiled about the records and aggregations of records already in the preserver's custody for the purposes of helping make valuation determinations during appraisals and helping facilitate accessioning of accruals during acquisition;
e. *information for arrangement*—that is, documentation compiled about acquired and accessioned records and their preservation for the purpose of arranging the preserved records of a given creator;
f. *information for description*—that is, documentation compiled about acquired and accessioned records and their preservation for the purpose of describing preserved records and creating descriptive instruments. Among other things, as stipulated in InterPARES Baseline Requirement B.2, such documentation should include explicit information about:
    - the date of the records' reproduction and the name of the responsible person;
    - the relationship between the records acquired from the creator and the copies produced by the preserver;
    - the impact of the reproduction process on their form, content, accessibility and use; and
    - an indication of those cases where a copy of a record is known not to fully and faithfully reproduce the elements expressing its identity and integrity;[53] and
g. *information for storage of preserved records*—that is, documentation compiled about preserved records and their elements and digital components for the purpose of facilitating their storage and long-term preservation.

3. *to update information on preservation activities* (A4.4.1.2), which involves recording information about actions taken to update digital components of records acquired for permanent preservation or their storage. Updating digital components and storage of digital records takes place over time and recording information about these actions is an important responsibility of preservers to be able to supply it to users wishing to have evidence to support the presumption of authenticity of the records. As envisioned in the model, this activity serves to compile metadata generated during storage maintenance activities (back-up, correction, update and refreshment activities) and update the preservation information system. This updated maintenance information will, in turn, be

[53] Ibid., 213.

used to periodically update the maintenance information provided in the records descriptions. Metadata associated with records aggregations or individual records are recorded, as appropriate, including:

- maintenance activity identification number(s) (i.e., backup, correction, update or refreshment identification number(s), as a mechanism for location of the record(s) and for linking to the relevant maintenance activity report(s)/metadata); and
- accession number(s).

Managing the Arrangement of Preserved Records (A4.4.2)

This activity provides overall control and co-ordination of arranging the records of a creator that have been identified as to their provenance and relationships according to the concepts and principles of archival arrangement.

Managing the Description of Preserved Records (A4.4.3)

This activity provides overall control and co-ordination of records description activities, including monitoring the preservation description system, describing preserved records and developing description instruments. Specifically, the component activities of managing the description of preserved records are:

1. *to monitor the performance of the preservation description system* (A4.4.3.1), which involves assessing the efficacy of the performance of the permanent preservation description system by analyzing reports on the operation of description management activities and, in response, issuing (1) activity directives for description activities and (2) reports on the performance of the description system for use in continued maintenance of the permanent preservation system;
2. *to describe the preserved records* (A4.4.3.2), which involves recording information about the nature and make-up of individual records and/or records aggregates acquired for permanent preservation and about their juridical-administrative, provenancial, procedural, documentary and technological contexts, as well as information about any changes they have undergone since they were first created. As noted earlier, archival description is one of the key means for authenticating the records in the custody of the preserver and for perpetuating the administrative and documentary relationships of the records, so it is important that special care and attention is exercised in compiling, recording and capturing all the identity and integrity metadata needed to meet the requirements supporting the production of authentic copies of the creator's records.[54] Such metadata could include:
   a. Identity metadata (may be inherited from higher description levels, as appropriate)
      - transfer registration number (Note that the record(s) being described inherit the identity/integrity metadata recorded in the transfer registration register.)
      - accession number (Note that the record(s) being described inherit the identity/integrity metadata recorded in the accessions register.)
      - accrual number (as appropriate) (Note that the record(s) being described inherit the identity/integrity metadata recorded in the accruals register.)
      - parent unit number (as appropriate) (Note that the record(s) being described inherit the identity/integrity metadata recorded for the parent unit.)

[54] The model assumes the baseline requirements as listed in the InterPARES 1 Project's "Requirements for Assessing and Maintaining the Authenticity of Electronic Records" (Authenticity Task Force, "Appendix 2," op. cit., especially 212–219).

b. Integrity metadata (if not evident in the metadata inherited from the transfer, accessions and/or accruals registers or from the parent unit)
    - indication of the original state of the record(s) when received (Note that state in this context is characterized in relation to the information for preservation carried forward from the appraisal process.);
    - indication of the security and control procedures used for records transfer, maintenance and reproduction activities;
    - indication of the current state of the records (Note that state in this context is characterized in relation to the updated information for preservation issuing from the processes of correcting, updating and/or refreshing digital components or storage.);
    - maintenance activity identification number(s) (as appropriate)—provides a link to information about any maintenance actions applied to the record(s) (e.g., correcting, updating, refreshing), and the impact of these actions on the form, format, authenticity, etc., of the record(s);
    - indication of any access restriction(s) related to copyright, privacy, etc.;
    - indication of the digital rights that apply to the record(s) being described, including:
        - name of the person(s) holding the rights;
        - terms and condition of the rights, including jurisdiction, duration, pertaining to which records, etc.; and
        - rights document number (e.g., deed of gift, contract, etc.);
    - location of the record(s) in storage;
    - date of the description;
    - name of the person responsible for the description; and
    - indication of the description rules used.

3. *to develop description instruments* (A4.4.3.3), which involves preparing tools that provide intellectual and physical control over the records in the preservation system, such as guides, inventories, indexes and repository locators.

Managing the Storage of Preserved Records (A4.4.4)

This activity provides overall control and co-ordination of the permanent preservation storage system and the records stored in the system by placing the records in storage, maintaining their digital components and monitoring the performance of the storage system. Specifically, the component activities of managing the storage of preserved records are:

1. *to monitor the performance of the preservation storage system* (A4.4.4.1), which involves assessing the efficacy of the performance of the permanent preservation storage system by analyzing reports on the operation of storage activities and, in response, issuing (1) activity directives for storage activities and (2) reports on the performance of the storage system for use in continued maintenance of the permanent preservation system;
2. *to place preserved records in storage* (A4.4.4.2), which involves placing the digital components of preserved records and their metadata into storage in accordance with the procedures for maintaining authentic copies of records and the actions prescribed by the preservation storage system strategies, rules and procedures and activity directives. Integrity metadata captured for this activity could include:
    - indication of the original state (e.g., file format) of the record(s) prior to storage;

- indication of any modification(s) made to the record(s) in preparation for storage;
- indication of the state of the record(s) after the modification(s) (e.g., impact on form, format, authenticity, etc.);
- indication of the reason/authorization for the modification(s) (e.g., through reference to the relevant section of the preservation storage system strategy);
- date/time of any modification(s);
- name of the person responsible for the modification(s);
- name of person responsible for placing the record(s) in storage;
- date/time the record(s) was/were placed in storage; and
- location of the record(s) in storage.

3. *to maintain the records in the permanent preservation storage system* (A4.4.4.3), which involves monitoring the storage of preserved records and their digital components, periodically backing-up the permanent preservation storage system and, as necessary, correcting problems with and updating the digital components, and/or refreshing the storage media to ensure that the records in the system remain accessible, legible and intelligible over time. In particular, the component activities of maintaining the storage of preserved records are:
   a. *to monitor the preserved records in storage* (A4.4.3.1), which involves keeping track of the condition and maintenance requirements of preserved records—more specifically, their digital components and metadata—and the media on which they are stored in the permanent preservation storage system to identify storage that needs backing-up, digital components and metadata that need correcting or updating and media that need refreshing. Another key task of this activity is to issue reports on maintenance activities to the preservation information system;
   a. *to back-up the permanent preservation system* (A4.4.3.2), which involves routinely creating copies of the digital content in the preservation storage system for the purpose of recovery in the event of a disaster resulting in system failure or corruption, and recording information about the back-up activities. It is important to distinguish here between comprehensive *system backups* and localized *content backups*. System backups contain a copy of *all* the digital objects in the system, including the operating system, the software applications and all digital objects (i.e., digital components of the records and their metadata) in the system. A system backup provides the maximum level of recovery potential in the event of a disaster or system corruption. Content backups contain a copy of selected aggregations of the digital objects in the system and, therefore, only offer limited recovery potential, especially in cases where system applications become corrupted. Integrity metadata about these back-up activities, captured in activity reports could include:[55]
      - indication of the reason/authorization for the backup (e.g., through reference to the relevant section of the preservation storage system strategy);
      - indication of the type of backup (e.g., incremental, differential, full) ;
      - indication of the extent or content of the backup (e.g., full system, selected groups of records, etc.);
      - name of the person creating the backup;

[55] As per InterPARES 1 Baseline Requirement B.1 - Controls over Records Transfer, Maintenance, and Reproduction (Authenticity Task Force, "Appendix 2," op. cit., 213).

- date/time of the backup;
- indication of the software application (including version number) used to create the backup;
- location of the backup; and
- backup identification number.

b. *to correct problems with the preserved records in storage* (A4.4.3.2), which involves taking the actions prescribed by the preservation storage system strategies, rules and procedures and activity directives, in accordance with the procedures for maintaining authentic copies of records, to identify and eliminate problems in storage to ensure that the records remain accessible, legible and intelligible over time; and recording information about the correction activities and the corrected digital components. Integrity metadata about this activity, captured either as metadata attached to the records or in activity reports, could include:[56]
   - indication of the original state (e.g., file format) of the record(s) prior to correction;
   - indication of the correction processe(s) used;
   - indication of the state of the record(s) after correction (e.g., impact on form, format, authenticity, etc.);
   - indication of the reason/authorization for the correction (e.g., through reference to the relevant section of the preservation storage system strategy);
   - name of the person responsible for the correction;
   - date/time of the correction; and
   - correction identification number, as a mechanism for location of the record(s) and linked to an accession number(s).

b. *to update the preserved records in storage* (A4.4.3.4), which involves carrying out conversion actions (such as by migration, standardization or transformation to persistent form) on the digital components of the preserved records in storage in accordance with the procedures for maintaining authentic copies of records and the actions prescribed by the preservation storage system strategies, rules and procedures and activity directives, to ensure that the records remain accessible, legible and intelligible over time; and recording information about the updating activities and the updated digital components. Typical conversion activities might include migration, standardization or transformation to persistent form. Integrity metadata related to this activity would be similar to those noted above for correction activities;[57] and

c. *to refresh the storage media for the preserved records in storage* (A4.4.3.5), which involves copying or transferring the digital components of preserved records in storage from one medium to another—or otherwise ensuring that the storage medium remains sound—in accordance with the procedures for maintaining authentic copies of records and the actions prescribed by the preservation storage system strategies, rules and procedures and activity directives; and recording information about the refreshment activities and about any impact to the digital components on the refreshed media. Again, integrity metadata related to this activity would be similar to those noted above for correction activities.[58]

[56] Ibid.
[57] Ibid.
[58] Ibid.

Outputting Records (A4.5)

This final preservation system activity involves facilitating the discovery of preserved records and/or information about records in the permanent preservation system, managing access and retrieval requests for the records and/or information—including presenting the records and/or information to users or packaging the records/information for issuing to users—and monitoring the performance of the permanent preservation retrieval and access systems.

Monitoring the Performance of the Permanent Preservation Access System (A4.5.1)

This activity involves assessing the efficacy of the performance of the preservation access system by analyzing reports on the operation of preservation access activities and, in response, issuing (1) directives for access activities and (2) reports on the performance of the access system for use in continued maintenance of the permanent preservation system.

Facilitating the Discovery of Preserved Records and/or Information (A4.5.2)

This activity provides authorized internal and external users with mediated access to and, as necessary, assistance in the use of, record descriptions, description instruments and any other tools and resources provided by the preserver to support querying and searching for information, records and/or records aggregates in the permanent preservation system.

Managing Requests for Preserved Records and/or Information (A4.5.3)

This activity provides overall control and co-ordination of internal and external requests for access to preserved records and/or information about the records by processing access requests, retrieving digital components for requested records and/or information, verifying retrieved components and information and providing access to retrieved records and/or information. Specifically, the component activities of managing requests for preserved records and/or information are:

1. *to process requests for preserved records and/or information* (A4.5.3.1), which involves the following four activities:
    a. *to register preservation access requests* (A4.5.3.1.1), meaning to record information about the request to register the circumstances of its occurrence and to issue notifications of receipt to the persons requesting the records. In cases where requests include insufficient information for registration purposes (e.g., incomplete information about the records or information being requested), the notification of receipt would include a request to the user to update the request. Specifically, this activity involves capturing some or all of the following metadata in an access register or similar instrument:
        - name of the person requesting the records/information;
        - name of the person for whom the request is being made (if different than the requestor);
        - access privileges of the requestor (as appropriate);
        - indication of the records and/or information requested;
        - date and time the request was received/registered;
        - name of the person registering the request;
        - access request registration number (as assigned by the preserver);
        - name of the person to whom a notification of receipt of request was issued;
        - indication of additional information required to register request (if necessary)

- name of the person who issued the receipt notification; and
- date and time the receipt notification was sent.

b. *to retrieve information to process preservation access requests* (A4.5.3.1.2), meaning to gather the information, from record descriptions and other descriptive instruments and from information about the preserved records in storage and their digital components, that is needed to process access requests and translate them into retrieval requests that can be processed by the preservation information and storage systems;

c. *to generate preservation retrieval requests* (A4.5.3.1.3), meaning to translate the access requests into requests that can be processed by the permanent preservation storage and information systems for the purpose of retrieving the exact digital components and/or information required to fulfil the access requests;

d. *to generate preservation request specifications* (A4.5.3.1.4), meaning to issue instructions to the preservation retrieval and access systems on how to fulfil retrieval and access requests based on analyses of the requests and the request processing information in relation to the preservation retrieval and access systems' strategies, rules and procedures (including procedures for maintaining authentic copies of records) and access privileges.

2. *to retrieve the requested records and/or information* (A4.5.3.2), which involves outputting copies of the digital components of the requested records, information about the digital components and rendering and/or content information about the records retrieved from storage in the permanent preservation system in response to retrieval requests for the components and/or information and in accordance with any request specifications;
3. *to verify the retrieved records and/or information* (A4.5.3.3), which involves examining the digital components and/or information retrieved to determine whether all the components and information that were requested have been received and can be processed for output in accordance with the current preservation access strategies applicable to those records. In cases where digital components are encountered that need updating or correcting, they are redirected, along with information about the problems encountered, to the maintenance function of the permanent preservation storage system for further action. Moreover, for retrievals containing digital components that cannot be processed or that are incorrect or incomplete, an order to rectify the retrieval may be issued.

   a. If the completeness, accuracy and ability to process the retrieved components and information is verified, the following information is recorded as metadata in a retrieval register or similar instrument:
      - date/time the retrieval was accepted as verified;
      - indication of the measures used to verify the retrieval;
      - name of the person verifying the retrieval;
      - retrieval verification registration number; and
      - retrieval request registration number.

   b. If the retrieval cannot be verified, the following information is recorded as metadata in the retrieval register:
      - date/time the retrieval was rejected;
      - name of the person rejecting the retrieval;
      - name of the person to whom an order to rectify the retrieval was issued;
      - name of the person who issued the order;

- date and time the order was sent;
- indication of the measures used to assess the retrieval;
- indication of the reason(s) for the rejection;
- retrieval rejection registration number; and
- retrieval request registration number.

4. *to provide access to retrieved preserved records and/or information* (A4.5.3.4), which involves providing users with access to copies of preserved records and/or information about the records. For certain requests, such as those under copyright and privacy laws, it may be necessary to document and keep information about the records and/or information issued in response to requests, to whom the records and/or information were issued and when. In cases where redacted records are issued, a copy of the redacted record should be created and its existence properly documented with the appropriate metadata. Unsuccessful attempts to provide access to the records and/or information trigger the creation of a "notification of rejection of preservation access request" that is issued to the requestor, a record copy of which is also retained by the preserver. The component activities of providing access to retrieved preserved records and/or information are:
   a. *to reconstitute the preserved records and/or information* (A4.5.3.4.1), meaning to link or assemble all the verified digital components of the requested records and/or information for the purpose of reproducing and manifesting (presenting) to the user copies of the records and/or information in authentic form and, if necessary, to redact information to meet privacy and/or copyright requirements. Integrity metadata captured for this activity could include:
      - indication of any problems encountered in reconstituting the records and/or information in authentic form;
      - indication of required maintenance action(s);
      - indication of any redaction for privacy or copyright reasons;
      - indication of the reason/authorization for the redaction;
      - date of the redaction;
      - name of the person responsible for handling/executing the redaction; and
      - registration number of the record copy of the redacted record issued to the user;
   b. *to manifest preserved records and/or information* (A4.5.3.4.2), meaning to present to the user copies of the reconstituted requested records and/or information about the records with the appropriate extrinsic form and, in the case of records aggregates, with information about their relationships to one another (archival bond). This activity also involves providing users with a Certificate of Authenticity, if requested. Regarding metadata, this activity results in the production of two sets: one set of integrity metadata for the preserver to document the activity and one set of identity and integrity metadata for the user. Metadata captured for this activity could include:

      Designated Preserver (integrity metadata)

      For requests that are fulfilled (in part or in whole)
      - indication of the record(s) and/or information presented;
      - indication of any redaction for privacy or copyright reasons (as appropriate);
      - indication of a Certificate of Authenticity, if issued;
      - indication of the means by which the records were authenticated
      - name of the person to whom the record(s) and/or information were presented;

- date when the record(s) and/or information were presented;
- name of the person responsible for handling/effecting the access request;
- indication of the state or condition of the record(s) and/or information at time the request was fulfilled (including, especially, an indication of instances where a copy of a presented record is known not to fully and faithfully reproduce the elements expressing its identity and integrity);[59] and
- indication of any problems encountered in manifesting the records and/or information in authentic form.

For requests that cannot be fulfilled (in part or in whole)

- indication of why the request cannot be fulfilled;
- name of the person responsible for determining that the request cannot be fulfilled;
- indication that a notification of rejection was sent;
- name of the person to whom the rejection notification was sent;
- name of the person responsible for issuing the rejection notification; and
- date/time the rejection notification was sent.

Records User

Identity metadata

- name(s) of the person(s) concurring in formation of the record(s);
- name(s) of action or matter;
- date(s) of creation and transmission of the record(s);
- expression of archival bond; and
- indication of any attachments.

Integrity metadata (as necessary)

- indication of access privileges used to control preservation of the presented record(s);[60]
- indication of protective procedures used to prevent corruption of the presented record(s);[61]
- indication of protective procedures used to guarantee the continuing identity and integrity of the presented records against media deterioration and across technological change;[62]
- indication of the means by which the presented record(s) was/were authenticated;
- indication of instances where a copy of a presented record is known not to fully and faithfully reproduce the elements expressing its identity and integrity;
- indication of any redaction for privacy or copyright reasons;
- indication of the reason/authorization for the redaction;
- date of the redaction;
- name of the person responsible for handling/executing the redaction;
- date when the requested record(s) and/or information were presented; and
- name of the person responsible for handling/executing the access request.

[59] As per InterPARES 1 Baseline Requirement B.2.d (Authenticity Task Force, "Appendix 2," op. cit., 213.).
[60] As per InterPARES 1 Baseline Requirement B.1.b (Ibid.).
[61] As per InterPARES 1 Baseline Requirements B.1 - Controls over Records Transfer, Maintenance, and Reproduction and B.2 - Documentation of Reproduction Process and its Effects (Ibid.).
[62] Ibid.

c. *to package preserved records and/or information for output* (A4.5.3.4.3), meaning to combine the digital components of the requested records and/or information with instructions on how to reconstitute and manifest the records or information with the appropriate extrinsic form. Regarding metadata, this activity results in the production of the same two sets of metadata outlined above, with the exception of the indication of a Certificate of Authenticity.

Business-driven Recordkeeping Model[63]

Introduction

This part of the Modeling Cross-domain report:

- explains the principles and concepts that underpin the BDR model;
- explains the model itself;
- positions the model in relation to the business activities of creating organizations;
- positions the model in relation to current archival thinking; and
- suggests possible approaches to implementation of the model.

Relation to other research

This modeling effort was influenced and informed by related work undertaken by the Modeling Cross-domain as well as by work undertaken by other InterPARES 2 research groups. This work includes the development of the Chain of Preservation (COP) model, the case studies conducted in the focus groups and the work undertaken by the Description Cross-domain. The model also incorporates key components of the model of the preservation function that was produced under InterPARES 1.[64]

Beyond InterPARES, the model benefited from the results of work conducted through other initiatives such as the digital preservation cluster of the Delos Project,[65] the Clever Metadata Project and Records Continuum research at Monash University,[66] the work on records management standards in ISO TC46/SC11 (standards: ISO 15489:2001, and ISO 23081-1:2006),[67] and the DIRKS methodology developed by the National Archives of Australia.[68] The Open Archive Information System reference model (OAIS, ISO 14721:2003),[69] which served as the basis for the preservation function model for InterPARES 1, was also used as the basis for the model.

The model is designed to be technology independent. The preservation of digital information is heavily dependent on technology that can change rapidly through time. New approaches to managing records, including the application of grid technology, the use of registries, the emergence of service oriented architectures, the development of enhanced approaches to exploiting Web-based technologies and the maturation and adoption of related standards in business and technical domains, are among a very few of the many trends that are having a significant impact on the way digital information is being managed and preserved. Every effort

[63] The BDR model and this report were produced with the assistance of Babak Hamidzadeh (Library of Congress), John McDonald (private consultant), Kenneth Hawkins (U.S. National Archives and Records Administration) and William T. Underwood (Georgia Tech Research Institute).

[64] See Kenneth Thibodeau et al., "Part Three – Trusting to Time: Preserving Authentic Records in the Long Term: Preservation Task Force Report," in Duranti, *Long-term Preservation* , op. cit., 99–116. Online reprint available at http://www.interpares.org/book/interpares_book_f_part3.pdf).

[65] See http://www.delos.info/index.php?option=com_content&task=view&id=25&Itemid=51.

[66] See http://www.sims.monash.edu.au/research/rcrg/research/crm.

[67] See International Organization for Standardization, ISO 15489-1:2001 - Information and documentation—Records management—Part 1: General; and ISO 23081-1:2006 - Information and documentation—Records management processes—Metadata for records—Part 1: Principles, 2.

[68] See http://www.records.nsw.gov.au/recordkeeping/dirks-manual_4226.asp.

[69] See International Organization for Standardization, ISO 14721: 2003 - Space data and information transfer systems—Open archival information system—Reference model.

was made to position the model at a high enough level to be independent of these influences and yet at a level where it is possible to develop implementation strategies that are meaningful and practical.

Scope and objectives

As discussed in the first section of this report, the COP model adopts the perspective of the archivist concerned about preserving the accessibility of records generated by the creating organization. The reference point for the model is the archivist looking into the "business" of a creating organization and identifying the records that are deemed necessary to preserve for internal business needs, or are likely to contribute to wider historical or societal objectives and interests.

In contrast, the perspective of the Business-driven Recordkeeping (BDR) model is on the organization addressing its own "business" within broader juridical, economic and cultural contexts, and the records generated by that business. The viewpoint includes both those records needed for current business and those that need to be retained and preserved for the longer term historical interests of society. The overall intent of the model is to illustrate the nexus between the needs and activities of the business of a given organization and the records generated by those needs and activities and kept by the organization.

More than addressing the business and records nexus of a given organization, however, the model also seeks to establish a generic framework that can be used by any organization for managing records as long as required, independent of the specific business context of a given organization, the organizational structure and the scale or size of the organization (e.g., an incorporated individual, a multi-national private sector organization, a government, etc.). Finally, and in keeping with the 'continuum' concepts developed in Australia, the model also accounts for the position of the organization (and its business) within the larger context of "society."

Given its perspective (i.e., the business of a given organization or organizations), this model will be of greatest interest not only to records managers but also, and most importantly, to program managers; that, is those managing the "business" of the organization, including those accountable for the fiscal and legal standing of the organization. By illustrating how the management of records is an integral component of the management of any business, the BDR model seeks to break down the barriers that often surround recordkeeping and other business processes. These barriers can be conceptual—such as the perspective that records management is something distinct and separate from the business of the organization—and social, between program managers and record managers and ultimately the archivists who are poised to secure those records requiring long-term preservation to meet societal and related historical requirements.

The specific objectives of the model are as follows:

1. To offer an integrated view of the business of an organization and recordkeeping.
2. To support parallel contexts and multiple views and perspectives (i.e. not only those within the business but also those who may be stakeholders, clients, partners, etc., as well as others, such as archivists, who may have interests in the records generated by the business).
3. To provide a framework for:
   - identifying the legal, juridical, ethical, business, organizational and archival requirements of specific business lines and their juridical contexts;

- illustrating the relationships (nexus/connection) and dependencies between business activities, the meta-process of capturing records (evidence) of these activities and the processes for managing the records themselves;
- integrating recordkeeping requirements in business activities such that the records required to provide evidence of the business activities can be captured and preserved in an appropriate manner; and
- managing (authentic and reliable) records throughout their existence and within different contexts of use and interpretation.

Given the focus of InterPARES 2 on artistic and scientific endeavours as well as the activities of government organizations, the model has been designed to be applicable to any organization regardless of sector.

Research issues and methodology

Research approach

The objective of the Modeling Cross-domain was initially to integrate the models developed in the UBC Project and the InterPARES 1 Project in one overarching and integrated model that would illustrate how records could be managed to ensure their authenticity and reliability through time. It soon became clear, however, that it would be difficult for such a model to account for the two very distinct perspectives that would need to be respected—that of the archivist and that of the creator (i.e., the "business perspective").

The perspective of the archivist is based on the traditional concept that positions the archivist as a trusted third party with responsibility for long-term preservation of the trustworthiness (the reliability, authenticity and accuracy) of records and having the dictates of that preservation in mind. The other two parties are the records creator and users of the records external to the records creator, both of whom have an interest in maintaining the integrity of records. These interests, it is supposed in COP model, are asserted through the controls or constraints on the model.

The perspective of the "business" and the focus of this narrative are based on viewing records creation, use and preservation through the lens of the creating organization. According to this perspective records are generated by business processes and activities in support of the goals and priorities of a given organization. However, the interests of the organization in managing authentic and reliable records in support of these goals and priorities are augmented by the interests of other stakeholders such as partners, clients, auditors and society in general.[70] According to this perspective, and as a basic assumption of the model described by this narrative, the context(s) of records will always be at the forefront of concern. That is, records will be shaped not only by the organization that has need of them to support various business goals, but also by the various stakeholders who have their own goals. Partners, clients, auditors and society in general all have a role to play in the shape, nature and characteristics of the records generated in a given business process or activity. The interests of society in the records will be explicitly

[70] The concepts that form the basis for the model and this narrative were derived from the work undertaken in Australia on the "records continuum." An explanation of the records continuum can be found in: Frank Upward (1996), "Structuring the Records Continuum, Part One: Post Custodial Principles and Properties," *Archives and Manuscripts* 24(2): 268–285. Online reprint available at http://www.sims.monash.edu.au/research/rcrg/publications/recordscontinuum/fupp1.html; and Frank Upward (1997), Structuring the Records Continuum, Part Two: Structuration Theory and Recordkeeping," *Archives and Manuscripts* 25(1): 10–35. Online reprint available at http://www.sims.monash.edu.au/research/rcrg/publications/recordscontinuum/fupp2.html. A brief summary of continuum concepts is provided in Appendix 16.

recognized and taken into account when making and keeping records within the organization. In this model the archivist is actually an active player. The ability to be the archivist standing apart from the records creation, use and preservation processes is severely challenged in this model where it is expected that, in the ideal scenario, the interests of the archivist (i.e., on behalf of society or some other interest group) are accounted for and addressed within the same context as the interests of all other stakeholders.

Based on this perspective the actions that records managers and archivists (indeed anyone who is a stakeholder) perform on records must be documented and captured as evidence if the full "story" about the records is to be told. In this way, any future user will be able to know not only about the original transactions the records are accounting for, but also how the records themselves were managed through time. This is why the model addresses two levels of processes. The first level is the actual business process and the second level is the (meta-) processes involved in managing the evidence about the business process as reflected in the records. That is, one viewpoint focuses on the records of business transactions while the other looks at the records of the activities, processes and mandates integral to managing the records of those transactions. The advantage of incorporating both views is that a single comprehensive and highly integrated view can emerge based upon an explicit separation of the concerns of all stakeholders. In this way, business and archival decisions pertaining to what should be kept, how much should be kept, how long it should be kept, how it should be kept and why it should be kept, are informed by the widest range of interests.

Conceptual issues

Within the records continuum, the creation of records—that is, what records need to be made within the business process and why—is something that is seen as part of managing records. It should be determined as much as possible in the design of computer systems and software applications to enable required functions to be present in operational systems. From a business perspective it should be clear how the business process will be documented and why. To achieve this requires an understanding of the business, juridical and societal requirements in the given business context, which will also help determine what types of documents need to be created and with what technologies. This also touches on the question, when do records come into existence? Is it at the moment that they are registered in the records system, or is it after they already are created in the business process and have played a role in this process? From the records continuum point of view it will be the moment they are active in a transaction. This is part of the appraisal process, which is concerned with deciding (1) what records should be created to document a business activity and (2) how long those records should be retained.

The specific characteristics of digital records, especially the difficulty of preserving them across ever-new generations of information technology, make retaining them a challenge. Every transition to a new generation may entail loss of information and in some cases even loss of essential information that will impact on the authenticity of the records. Preservation becomes in this way part of the appraisal process too, because with each migration, conversion and even emulation an assessment has to take place to identify the best possible preservation strategy for a certain type of record. In this perspective, preservation is part of the process of maintaining records over time, which begins at the moment of designing the technological infrastructure and systems that will create, manage and maintain the records.

Digital preservation in this document is thus seen as the specific process of maintaining digital records during and across different generations of technology over time, irrespective where they reside.

The modeling techniques

It is impossible to model reality in all its aspects, relationships and expressions, not only because of the limitations of the human mind but more specifically because of the limitations of modeling itself. Modeling of reality in all its aspects provides a kind of meta-reality as incomprehensible as reality itself. This is why the intent of modeling per se is to take a "slice of reality" and to express certain aspects of that reality (or an object from that reality) graphically to understand how given objects are operating and interrelating. By virtue of its characteristics, a model seeks to reduce the complexity of reality to promote understanding that will enable informed actions and decisions to be taken.

The Preservation Task Force of InterPARES 1 developed a functional model of the process of preserving authentic digital records following the Integrated Definition (IDEF) method prescribed by the InterPARES International Team.[71]

> In IDEF0, "A function model is a structured representation of the functions, activities or processes within the modelled system or subject area."[72] An IDEF(0) model includes activities and entities. An activity is depicted as a box whose name indicates the nature of the activity. An entity either goes into or comes out of a process (activity). Three types of entities go into a process: inputs (I) that are transformed or consumed in the process, controls (C) that govern its execution, and the mechanisms (M) needed to carry it out. Only one type of entity comes out of a process: the outputs (O) that are produced by acting on the inputs under conditions and constraints imposed by the controls. In IDEF(0) diagrams, the four types of entities are always depicted as arrows in the following arrangement: Inputs enter a process box at the left side. Controls enter at the top. Outputs exit from the right, and mechanisms enter at the bottom. Given this invariant order, the entity arrows are collectively referred to as ICOMs.[73]
>
> In IDEF(0) diagrams, there are two basic icons: boxes are used to represent activities or processes and arrows represent ICOMs. In IDEF(0), a process may be decomposed into its sub processes. This is depicted by creating a new, child diagram in which the parent process box becomes the outer boundary of the diagram and the sub-processes are depicted as boxes within that diagram. All ICOMs connected to a box at a higher level are shown entering or exiting at the corresponding edge of the decomposition diagram. Successive decompositions can be delineated to achieve whatever level of precision or clarity is desired. Such successive decompositions constitute a decomposition hierarchy. All IDEF(0) models start at the highest level, labelled "A0," showing only one process box, which is the function being described taken as a whole, and the ICOMs that enter the function from the outside and that are output from the function. This simple notation provides a systematic and highly coherent method for describing a process to whatever degree of granularity is needed.

An example of such a diagram is shown in Figure 1 below.

[71] The following section has been quoted from Thibodeau et al., "Preservation Task Force Report," op. cit., 103.
[72] National Institute of Standards and Technology, *Draft Integration Definition for Function Modeling*, op. cit.
[73] Robert P. Hanrahan (1995), "The IDEF Process Modeling Methodology." Available at http://www.stsc.hill.af.mil/crosstalk/1995/06/IDEF.asp.

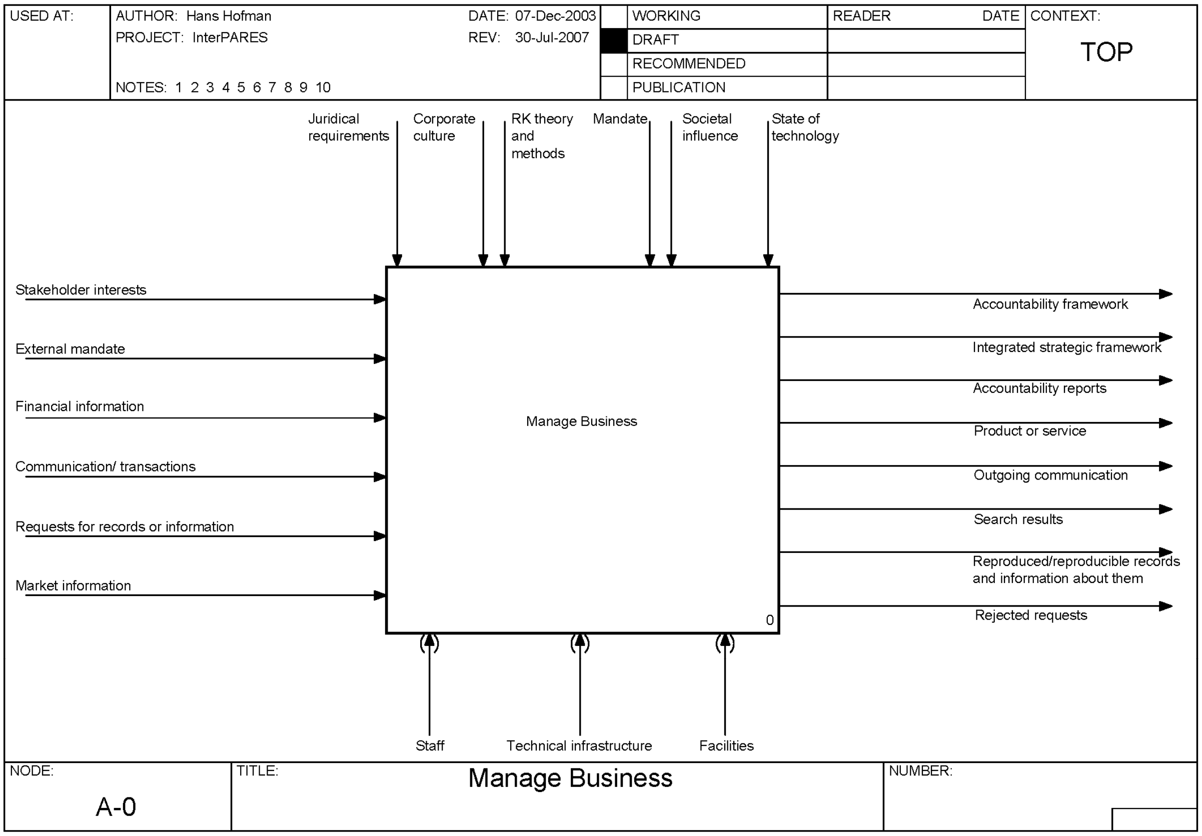


Figure 2. Highest level diagram of BDR Model in IDEF0 (A-0)

Although IDEF0 was viewed as adequate for the purposes of the development of the model, it was determined that certain adjustments would be needed to ensure that the model that emerged would be relevant and effective and, above all, capable of being implemented. To ensure that the modeling work captured essential aspects of the record continuum and its management, UML (Unified Modeling Language) class diagrams were also developed at a high level. UML class diagrams offered an object-oriented view that complemented the function-oriented view of IDEF0. UML uses a standardized graphical notation to represent classes of objects (including digital objects or artefacts but also the organizational and societal elements, instruments and constructs), along with their supporting information, expressing their interrelationships, characteristics and operations or behaviours.

> First and foremost, UML is a language. There are rules governing how things can be combined to form expressions. The idea is that UML doesn't just give you a way to draw diagrams, but it gives you a way to express concepts and relationships.[74]

The advantage of using UML class diagrams is that they allow identifying the classes of digital objects to be taken into account with their identity and other (meta)data elements, as well

[74] Dan Pilone, *UML 2.0 Pocket Reference* (O'Reilly Media, 2006), 3. UML is a specification of the industry consortium Object Management Group (OMG) and also is available from ISO as ISO/IEC 19501. More information about UML can be found at http://www.omg.org.

as the operations that can be performed on those (data) objects and the cardinality of their relationship with other objects.

Both the IDEF0 and the UML class diagrams are needed, since they each express important concepts in managing records from different perspectives. The IDEF0 model shows what processes, functionalities and activities are needed to support recordkeeping and management. It also shows how data and information are created, transformed and consumed by different functions. The UML class diagrams represent digital information objects that are central to managing records (e.g., records, retention schedules, etc.) and their composite structure, when relevant. The class diagrams also represent important attributes of the digital objects and show how the information objects relate to one another. It is possible to draw relationships between the IDEF0 model and the UML class diagrams. For example, data elements that comprise ICOMS in the IDEF0 model may appear as objects in the class diagrams with their attributes and relationships elaborated in the class diagrams.

In addition to employing UML class diagrams and recognizing the need to support multiple views, it was decided to adopt a liberal approach to the use of IDEF0 by stretching it in time and space (domain) as well as in levels of aggregation, thus making it scalable and more dynamic in its application.

The modeling work was undertaken by a small number of specialists comprising an archivist, an archival program analyst, a scientist and a computer scientist. Research assistants from the University of British Columbia supported the work by undertaking research on business models and definitions. As a result of time limitations it was not possible to conduct a "walkthrough" of the model, which is why the model still requires testing and validation and should be viewed as an intermediate result of the overall InterPARES 2 Project.

Applying the model

Principles that underpin the model

In understanding the model, there are some important principles that need to be taken into account. These principles, which serve as the basis for the two different expressions of the model, are as follows:

1. the model must be applicable to all types and sizes of organizations, ranging from an individual to a multinational;
2. the model must be applicable through time and through multiple dimensions and thus expressing more than one view and perspective; and
3. the model must be technology independent.

The model must also address the two dimensions of preservation: time and space. Through time, the authenticity, completeness, meaningfulness and usability of records needs to be maintained, despite changing cultural contexts and changing technologies.

Across space, different organizations representing multiple domains may be involved in the management of records both contemporaneously and sequentially through time. For instance, records created and used in a record creating organization may also be of interest to third parties. Across space, for instance, they may be of interest to citizens, audit authorities, tax revenue office, stakeholders and other parties depending on the business context. Through time, they may be managed by subsequent agents that have a responsibility for their preservation.

The model needs to reflect the most complicated situations and be applicable to organizations having complex mandates, organizational structures, business processes, etc. At the same time, however, the model must also be capable of being applied to the very smallest unit of "organization"—the individual, one-person enterprise. Needless to say, in the case of the one-person enterprise, the detailed diagrams and other complex aspects of the model may not be necessary. Regardless, when used it will be necessary to identify clearly at what level of aggregation the model is to be used. Within a given organization, for instance, it will be important to for the organization to determine whether the model is to be used for the whole organization, thereby including all business processes of the organization, or for one specific business process.

The model is thus indifferent to underlying organizational structures. What it tries to express is the coherence of primary business activities and the meta-processes of recordkeeping, no matter what organizational structures may exist. The overall intent is to support a perspective where the management of the business and the management of the records of the business are coherent and harmonized.

Benefits of the model

By taking the organization and the organization's business as a starting point, the model will help organizations ensure that their recordkeeping policies and strategies are in direct line with and supportive of their business needs. By embedding recordkeeping within the business, organizations will be able to secure the following benefits (as expressed in ISO 15489:2001, clause 4):[75]

- conduct business in an orderly, efficient and accountable manner,
- deliver services in a consistent and equitable manner,
- support and document policy formation and managerial decision making,
- provide consistency, continuity and productivity in management and administration,
- facilitate the effective performance of activities throughout an organization,
- provide continuity in the event of a disaster,
- meet legislative and regulatory requirements including archival, audit and oversight activities,
- provide protection and support in litigation including the management of risks associated with the existence of, or lack of, evidence of organizational activity,
- protect the interests of the organization and the rights of employees, clients and present and future stakeholders,
- support and document current and future research and development activities, developments and achievements, as well as historical research,
- provide evidence of business, personal and cultural activity,
- establish business, personal and cultural identity, and
- maintain corporate, personal or collective memory.

The model is a tool that organizations can use to systematically analyze, design and begin to implement their recordkeeping framework such that it emerges as an integral component of the organization's business. Given that most organizations have at least some of the components of recordkeeping in place (i.e., no matter how modest these components might be, the organization's are not starting with a clean slate), the model can be used to audit the existing situation and assess the extent to which various approaches to designing the recordkeeping function may or may not be appropriate to the given business context it is serving. The results will enable decisions to be made concerning not only the design of the recordkeeping function

[75] International Organization for Standardization, ISO 15489-1:2001, op. cit., 4.

and paths to implementing a recordkeeping system, but also how identified gaps between the recordkeeping function and the business functions of the organization can be closed.

In this respect, the UML class diagrams of the model identify the main *information objects* that are relevant in recordkeeping, such as agents, business and recordkeeping activities and records. For each of these entities, information has to be captured and kept over time. These information objects may be further detailed in sub-classes and through their interrelationships. All will be described through their attributes and, to some extent, through the operations that can be performed on them. Thus, the class diagrams help identify what information has to be created, captured and maintained over time in a recordkeeping environment.[76]

In contrast, the IDEF0 component of the model focuses more on the *activities* that need to be undertaken to implement recordkeeping in an organization. Those activities regard the recordkeeping needs, risks, requirements and subsequent policies, strategies, procedures and instruments, as well as the actual recordkeeping activities that are needed and appropriate in the organization, given the business functions and activities for which it is responsible.

The two views (the information-focus UML class diagrams and the activity-focus IDEF0 diagrams) are complementary and provide a good framework for organizations to identify what has to be done to implement good recordkeeping.

The model provides a framework and tool to:

- support organizations to implement recordkeeping in a systematic and structured way taking into account the current situation and the capability of the organization to achieve, if needed, better recordkeeping policies and practices;
- ensure that business activities have proper recordkeeping processes and procedures implemented; and
- ensure recordkeeping processes are properly embedded in business activities.

Finally, a model of a prospective operational recordkeeping environment enables the organization to conceptualize, plan, acquire resources, communicate with both technical and business stakeholders and evaluate, make and, if necessary, revisit decisions about realizing the model's functions in systems or services.

Overview of the Business-driven Recordkeeping Model[77]

In this section, the BDR model as developed with the two different modeling techniques (IDEF0 and UML Class Diagrams) will be discussed, as far as possible, in parallel to provide the reader with the opportunity to compare the two different expressions.

Introduction

The first objective of the model is to identify and express the relationship between doing business activities and the information used and created in doing this, as well as the capture and management of that information as records. The second process is the management of both sets of activities.

[76] Guiding the class diagrams was also the draft technical specification ISO-23081-2 on metadata for managing records.
[77] The Business-driven Recordkeeping model's IDEF0 diagrams, together with the IDEF0 activity and arrow definitions and the UML class diagram definitions, are provided in Appendix 15.

In taking this perspective one can identify three types of processes: the actual business processes, the actual meta-process of managing records (recordkeeping) and the overarching higher level meta-process of managing (the interaction between) these two.

The different types of processes (or contexts) leads to three distinct views expressed in UML class diagrams:

- *the business view*—taking the (organizational) perspective of a business in its activities and creating records; the core item for business is the transaction. From this view recordkeeping is viewed as a function of business;
- *the recordkeeping view*—taking the perspective of capturing and managing records created in a business; the core item for recordkeeping is the record. From this view a business activity provides the context of the records; and
- *the framework view*—taking the perspective of the need to integrate both the business view and the recordkeeping view so that they are properly aligned. This reflects the layer that controls and amalgamates or integrates both previous views.

In principle, these three views occur in any business situation. Depending on the scale, the complexity of interaction between them will increase and will lead to differentiation in the processes at each level with inherent proper organizational structures. A one-person shop does not need to differentiate (organizationally) between all the management processes that will be closely knit together. In a multinational organization, whole departments may be dedicated to one specific process; such as the process of accounting or that of records management, for instance. One layer that may not be explicitly visible in the IDEF0 component of the model is the wider context within which all three of these views sit: the societal and cultural dimension. To make the continuum approach view complete, however, this dimension has to be added:

- *the societal view*—taking the perspective of the ambient contextual environment, needed to understand the external influences on a business and its recordkeeping practices. This layer has its own dynamics and is (usually) out of the control of the organization that runs a business.

It is reflected indirectly, as it should be translated into requirements and constraints that will govern business and recordkeeping business, and does not surface explicitly. For this reason, a UML class diagram has been developed to represent the societal view. In this diagram, the classes of business activity and agent should be seen at a higher level of abstraction/generality than in the business or recordkeeping views. The societal view is both the broader context of the inner dimensions and a perspective or entity in its own right that embeds the agents (organizations, individuals) of which it consists.

Each of the four views does not stand on its own, but has an overlap with the other three. As such, they are interconnected and represent a multidimensional view on recordkeeping in line with the records continuum concept.

The core diagram

Starting from the diagram that delineates the object or area of observation from its context, the IDEF0 methodology decomposes the functions and processes within that area into ever more detail, until a level is reached that does not need further decomposition.

As indicated, this model takes the view of an organization (ranging from an individual to a multinational company) as its starting point and identifies all processes needed to perform recordkeeping in relation to any type of business activity. The core diagram is shown in Figure 2 with the top level of the three most relevant areas of activity: managing the business framework (A1, including other subordinate frameworks such as the recordkeeping framework), carrying out the actual business activities (A2) and carrying out the activities for managing records (A3).

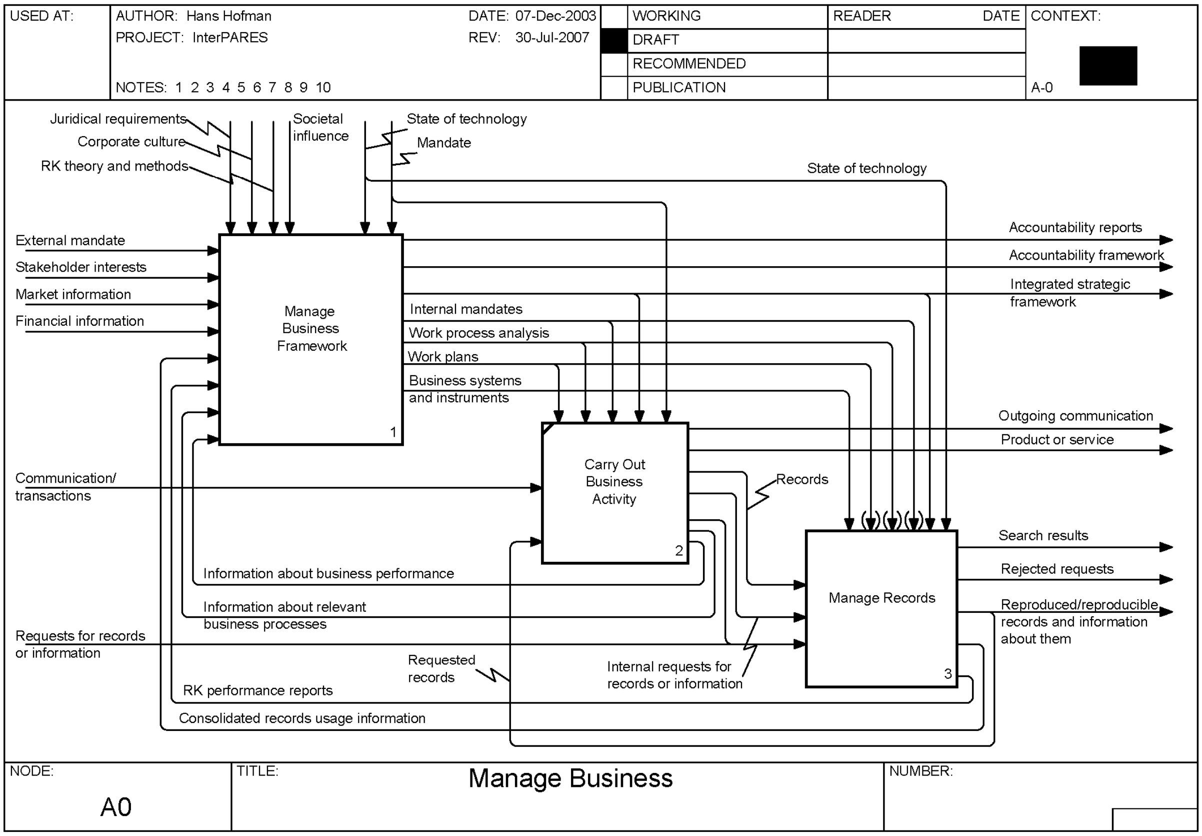


Figure 3. High level model of the business view (A0)

As indicated, in the IDEF0 methodology each of the boxes or processes in the A0 diagram is further decomposed into more detailed activities explaining what is needed to perform the high level process properly. In the domain of recordkeeping (box A3), the ISO 15489:2001 standard has served as a framework for identifying the processes involved, represented in activity boxes A1 and A3. For the box representing the carrying out of a business activity (box A2), this turned out to be more difficult for the domain of carrying out a business activity for two reasons. First, no appropriate generic reference model could be identified that could serve as a framework for it. Second, the alternative of using a real business activity would have provided an example of what could be expressed here, but at the same time would have limited the model to that one business activity. So it is left to the user of this model to take a relevant business activity or process as an example and decompose it into the steps that are needed to carry it out. When that is done, it will give "real world" meaning to the first and third boxes (A1 and A3), because they will be contextualized.

Thus, the model should be seen as an empty shell that can serve as an instrument for analyzing the business situation in relation to recordkeeping. The model becomes "alive" when box A2 is decomposed based on a concrete business activity. Doing so will also set the scope of the whole model.

Despite the above, it can be said that it is in carrying out a business activity that the three axes/dimensions of *transactionality*, *identity* and *evidentiality*, as identified in the records continuum model, come together. Agents (*identity*) carry out transactions (*transactionality*) that are documented in records (*evidentiality*). Each of these dimensions can occur at different levels of aggregation. The scope and complexity of the business will determine how many levels of aggregation may be necessary and how complicated the interrelationships are.

It has to be noted here that the evidentiality dimension also occurs at the two other levels of records: that is, the management of records themselves and the management of the business, both representing processes that have to be accounted for and subsequently documented, as well. As such, this represents the recordkeeping dimension.

Each of the three boxes in the A0 diagram of the IDEF0 model is also represented in UML class diagrams: the (Business and Recordkeeping) Framework view, the Business view and the Recordkeeping view, respectively. A fourth class diagram has been made representing the Societal view, which is comparable to the context diagram (A-0) of the IDEF0 model (see Figure 1).

Business activities

As mentioned above, box A2 in the IDEF0 diagram represents the actual business activities to be carried out. In principle, any actual business activity can be taken and subsequently be decomposed using the IDEF0 technique. Such decomposition will require the identification of the level of observation of the business function or activity and, subsequently, the analysis of the function or activity and how it is or should be performed. Such an analysis will identify not only how the business activity is carried out, but also by whom, when and why records will or should be created.

To complement the IDEF0 model, a generic, conceptual UML class diagram was made of the business activity (and transactions) and its relationship with records and record transactions and is given in Figure 3, below.

Four main clusters of objects can be identified:

1. the business transaction entity and the (business) activity to which it belongs or of which it is a part; in turn relationships are drawn to function, mandate, jurisdiction, business plan, etc.;
2. the agent entity responsible for the business activity, rights and associated recordkeeping;
3. the record entity and its metadata; and
4. in turn, these objects are governed or controlled by the classes of mandates and the business and recordkeeping frameworks.

The following class diagram does not reflect the fact that, in some cases, the record itself can be at the same time the business transaction.

The processes for managing policies and frameworks

Policies and strategies are essential for a successful recordkeeping function in an organization. Organizations will have mission statements based on the responsibilities they have for carrying out business functions and activities. Derived from that they will develop policies that determine what the objectives and goals are and what frameworks will be applicable for achieving these objectives and goals.

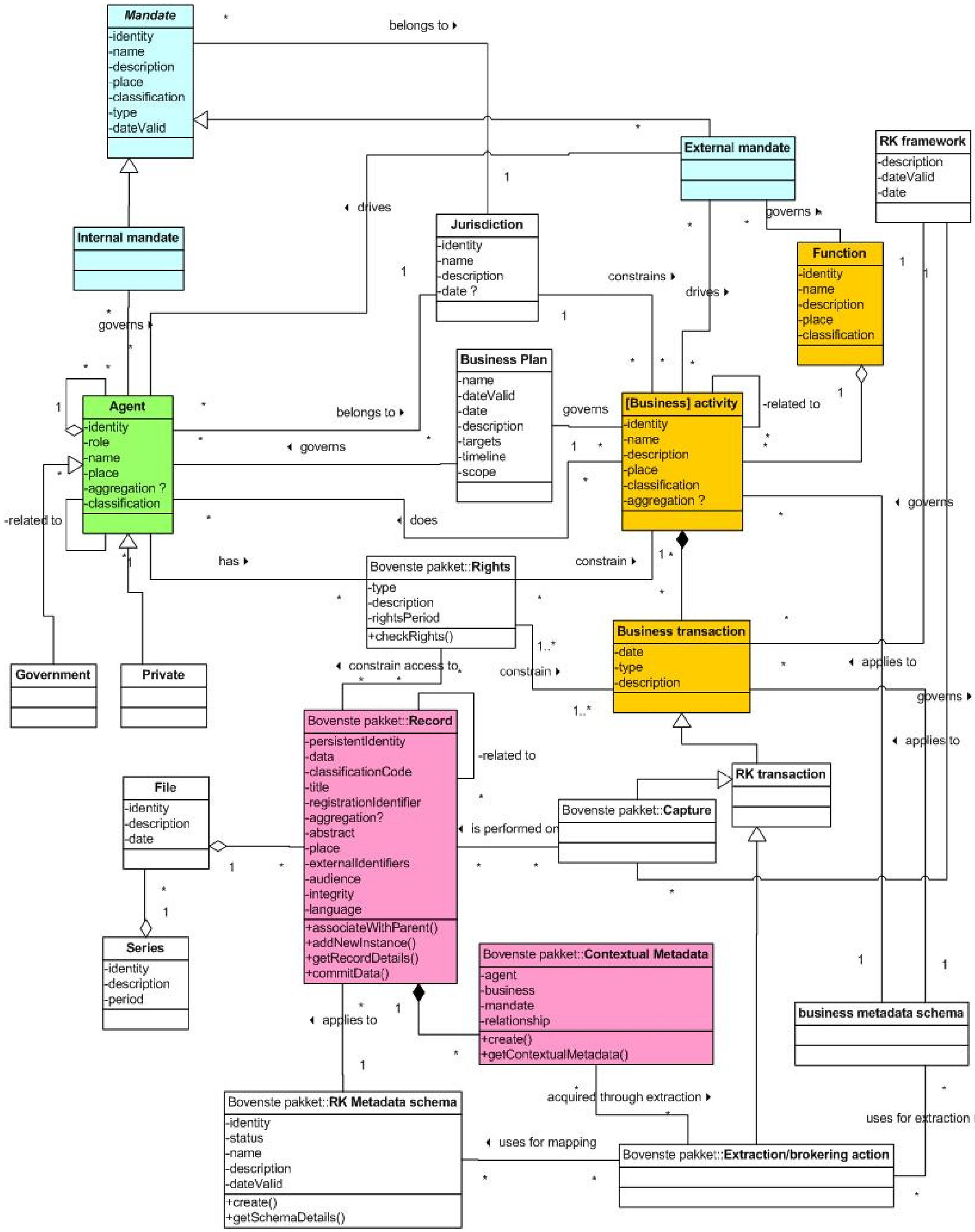


Figure 4. UML class diagram of business activity in relation to recordkeeping

In the following diagram (Figure 4), the high level functions are identified that will help organizations to develop the business framework and more detailed strategic plans. Policies serve two purposes: one internal, telling what has to be done according to what standards, rules and procedures, instruments and systems, and the other external, explaining to third parties under what policies or conditions the organization will perform its functions and obligations. An essential part of this activity should be a policy and framework for managing information and, more specifically, for managing the organizational records.

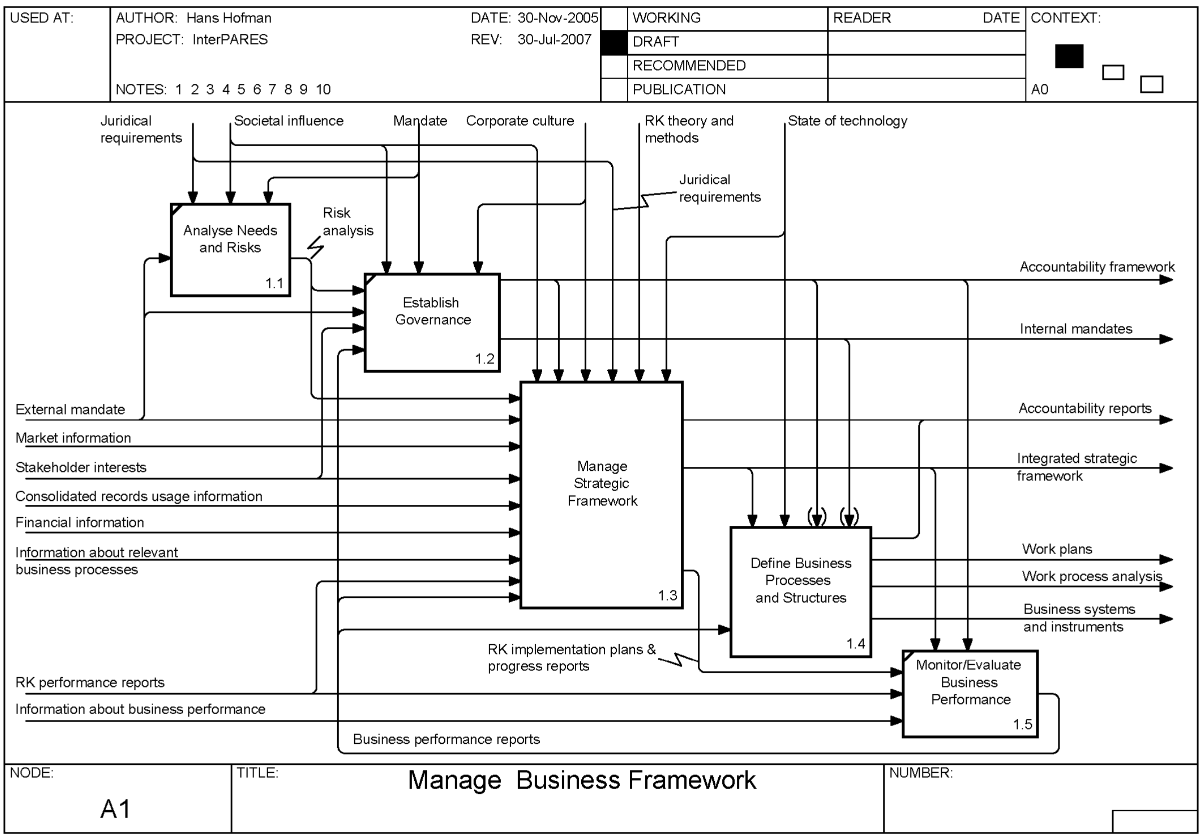


Figure 5. Overview of the main policy activities in managing a business (A1)

Such a framework and policy will be based upon an analysis of the organization's functions and juridical and societal context. This also will include an identification and analysis of risks and measures for how to mitigate those risks. The framework will guide all business and recordkeeping activities of the organization.

The processes involved are (IDEF0 model, A1; see Figure 4):

Analyse Needs and Risks (A1.1)

Establish the business needs for records (including how long they are required to be retained), identify the risks for the business and the organization that need to be mitigated by the proper creation and management of records, along with those risks that will be present if these records will not be properly created and managed, to identify the requirements for records management. This activity is controlled by legal, juridical and organizational or business requirements.

Establish Governance (A1.2)

Set the overall strategic direction of an organization by establishing the set of responsibilities and practices necessary to ensure that the organization is accountable for fulfilling its mandate, achieving its stakeholders' objectives, meeting the current societal, ethical and moral duties, managing its risks appropriately, using its resources responsibly and monitoring its performance effectively.

Manage Strategic Framework (A1.3)

Establish strategic plans outlining the organisation's current and future direction, priorities and resource allocation strategies, in line with its business needs and key stakeholder interests—as well as including the required mitigation of business risks identified—implement them within an overall strategic framework, monitor the performance and application of the established plans both for the business activities and the recordkeeping processes and, if necessary, subsequently adjust the plans to continue to meet business and key stakeholder needs and interests.

Define Business Processes (A1.4)

Define an organisation's business operational targets and outcomes, delegate and assign resources, develop business and workplans, and design and develop business work processes, necessary instruments and systems structures to effectively manage the organisation's resources and support its work processes.

Monitor and Evaluate Business Performance (A1.5)

Periodically assess the performance of the business processes in relation to the organisation's strategic framework and the accountability framework. Based on the monitoring, produce business performance reports to inform the organisation's appropriate management functions to confirm or revise the business strategic framework, or business processes and structures.

One of the purposes of the viewpoint expressed in the A1 view is to identify where the business framework meets the recordkeeping framework. This happens as part of the A1.3 activity (Manage Strategic Framework). One of the strategic plans that has to be developed is the recordkeeping framework (activity A1.3.4.3). This will be done in close connection with a risk management framework and the business framework.

Important activities in the process of developing a recordkeeping framework are apart from the development itself: namely, implementing it (A1.3.4.3.3) and evaluating its performance and adequacy (A1.3.4.3.4). Implementation is then guided by a readiness and capacity assessment to identify what is needed to transform the organization from the current situation to one that is better suited to address the recordkeeping requirements within the given business context.[78] This will not be a onetime activity, but will need to be repeated periodically, depending on the results of the monitoring process. This monitoring process will regard both the internal performance, such as whether goals are met or how many failures are occurring, and the external developments, such as change in technology or in legislation.

A good model for the whole management process is the "Deming Cycle," which distinguishes the processes of *plan* (define the policies and plans), *do* (implement the policies and plans), *check* (monitor and assess what is happening internally and externally), and *adapt* (decide on changes to improve).

[78] For a fuller overview of this approach, see http://www.tbs-sct.gc.ca/emf-cag/gcc-ec/ig-ge_e.asp.

The UML class diagram expresses a different view that focuses on the classes of information objects that are involved in setting policies and frameworks, their attributes, operations and relationships with one another.

In the diagram below (Figure 5), several clusters of classes can be distinguished:

1. the business activity cluster (gold) that is closely related to the business plan and the agent that is governed by a strategic framework;
2. the cluster for analyzing needs and risks (blue) leading to systematic and explicit identification requirements and an accountability framework;
3. the framework cluster (purple), including all types of policy frameworks necessary to do business; and
4. the classes mediating and interconnecting (light yellow) the different clusters such as the agent and the internal and external mandates.

As with the IDEF0 model, only a top level view is given here. For each aspect represented in Figure 5, new UML class diagrams at a lower level may be required that will represent a more refined and detailed view of underlying classes of activities and entities along with their attributes.

The recordkeeping processes

The third high level IDEF0 box (A3), which models the function of managing records, can be decomposed into four main processes (Figure 6) with the following definitions:

Capture Records (A3.1)

Based on rules established in the recordkeeping framework, the capture function identifies and brings under control the records that are created in the business activity and that need to be maintained. With the capture of those records, the required metadata are also captured/extracted to ensure the authenticity, usability, integrity and reliability of the records. The capture of metadata is done every time a record or aggregation of records is used in a business process. The capture process includes the registration and classification of the records as well as, if needed, the assignment of key words, so that the records are (uniquely) identifiable and searchable. The valid recordkeeping instruments will guide the registration and classification. Identification and information about the performance of this function are produced for evaluation purposes.

Maintain Records (A3.2)

This process is decomposed based upon the preservation function model as produced by the Preservation Task Force of the InterPARES 1 Project.

Following direction established in the preservation strategy as part of the recordkeeping framework for a given body of records selected for preservation, apply preservation method(s) targeted to that body of records to implement the preservation action plan for those records by maintaining the digital components of accessioned digital records, along with related information necessary to reproduce the records, certify their authenticity and enable correct interpretation of the records.

The maintain activity also carries out the disposition function, so that records are kept no longer than needed.

This maintenance activity enables the output, in response to a retrieval request, of the digital components of a record, along with information about that record or, if the request is only for information, the requested information. The "maintain" process also produces management or performance information that is used to evaluate execution of the "capture" function (A3.1). The process is carried out by persons responsible for preservation, using infrastructure technology.

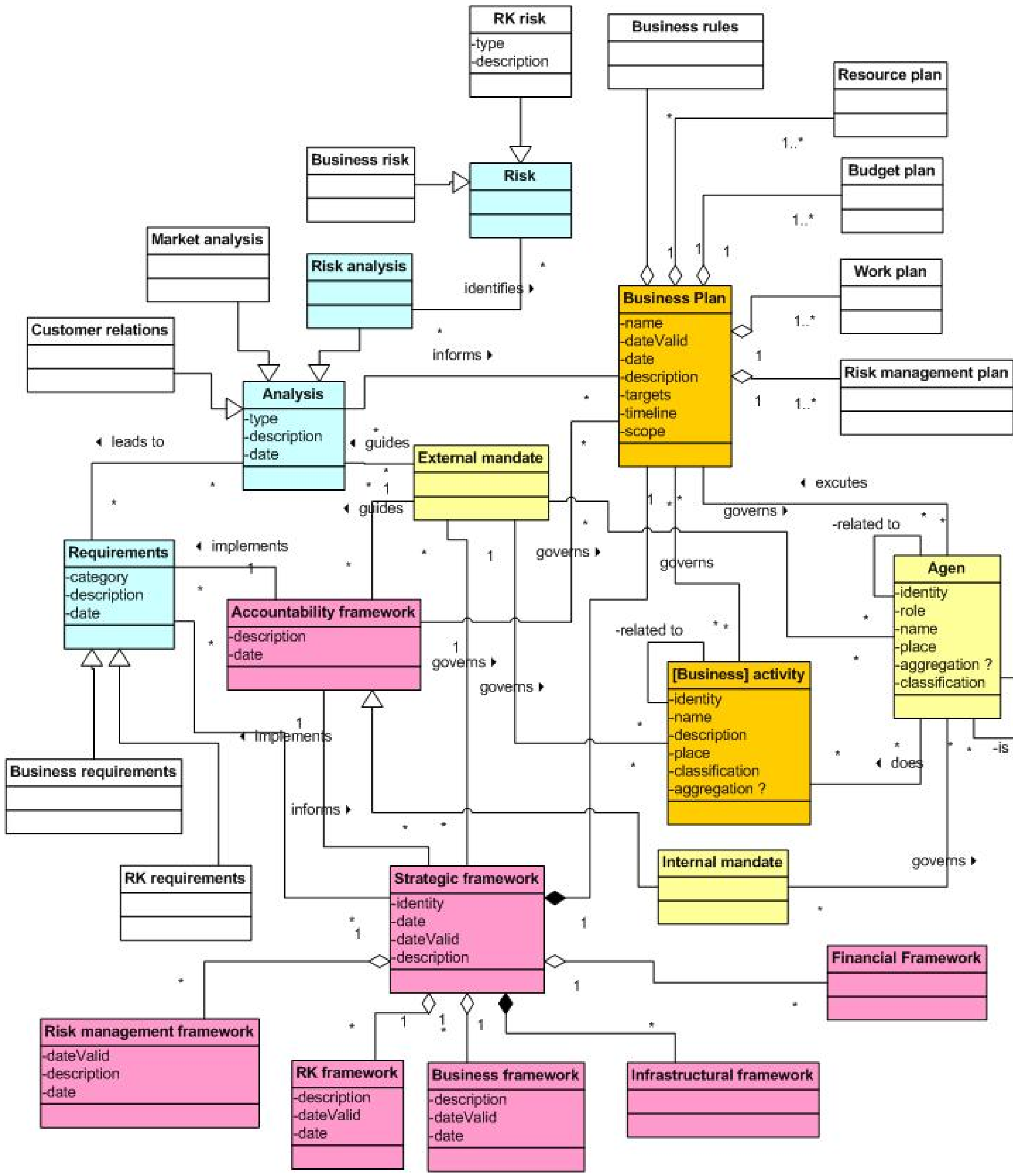


Figure 6. UML Class Diagram representing management of the different frameworks (including recordkeeping)

Facilitate Access (A3.3)

Governed by the access framework, support search facilities for users and, if successful, provide information about or provide access to reproduced (authentic) records or produce, if requested, a reproducible digital record; that is, the digital component(s) of the record along with instructions for producing an authentic copy of the record and information necessary to interpret the record as kept under the regime of the recordkeeping framework.

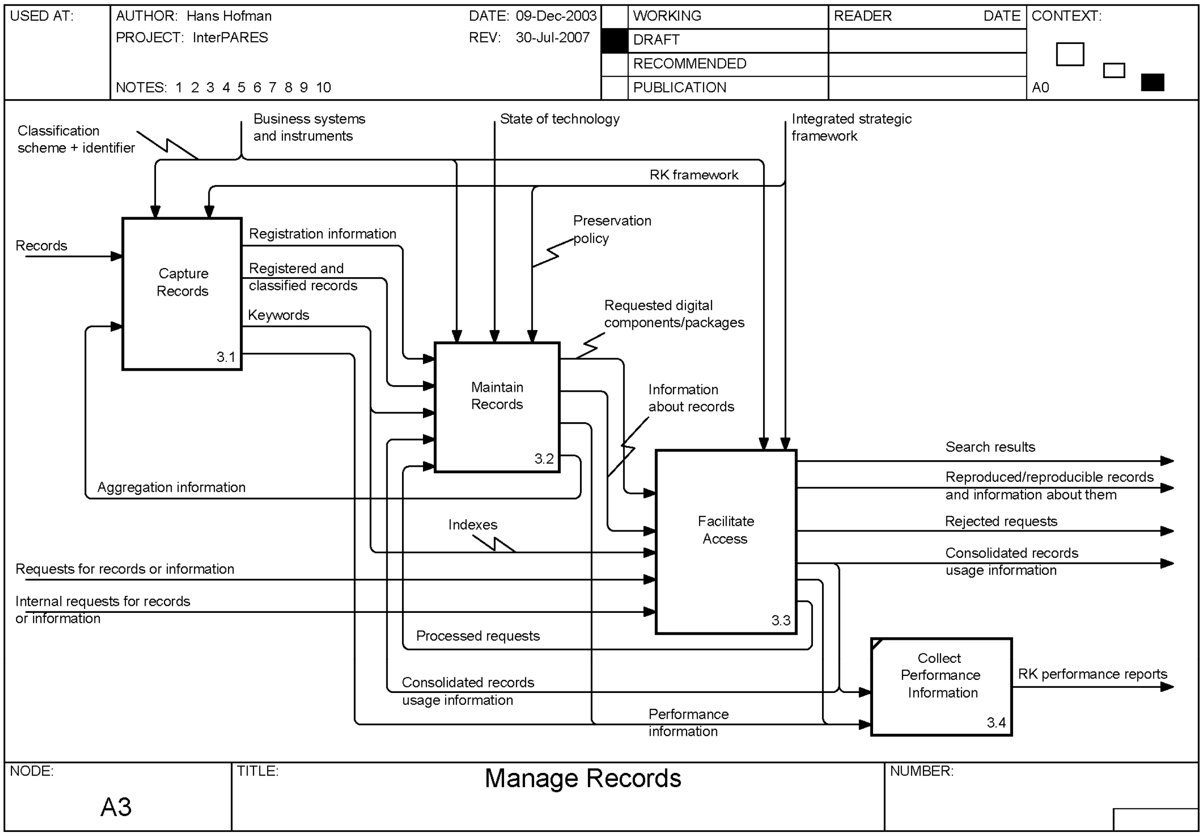


Figure 7. The Recordkeeping processes. IDEF0 - A3 diagram

Collect performance information (A3.4)

Synthesize and compile reports on the performance of the capture, maintain and facilitate access functions based on information continuously collected from these functions to inform the "evaluate recordkeeping framework performance and adequacy" function (A1.3.4.3.4). These reports may contain information about the applicability of policies, rules and methods, deviations from policies/rules and malfunctioning of systems, as well as suggestions for improvement. Other reports will be made with consolidated information about usage of records or aggregations.

The following UML class diagram (Figure 7) reflects a different expression of the recordkeeping function, drawing particular attention to the classes of objects involved in carrying out recordkeeping activities.

Those activities will be guided by a recordkeeping framework as established by the recordkeeping management function (see Figure 5). The diagram shows what transactions can be performed on the records (at all levels of aggregation), and the accompanying metadata, to maintain them as long as required. At the same time it identifies what metadata should be captured about the records and recordkeeping instruments and how these transactions should be documented. The classes of information objects and their attributes reflect, as such, a metadata schema needed to perform the recordkeeping function.

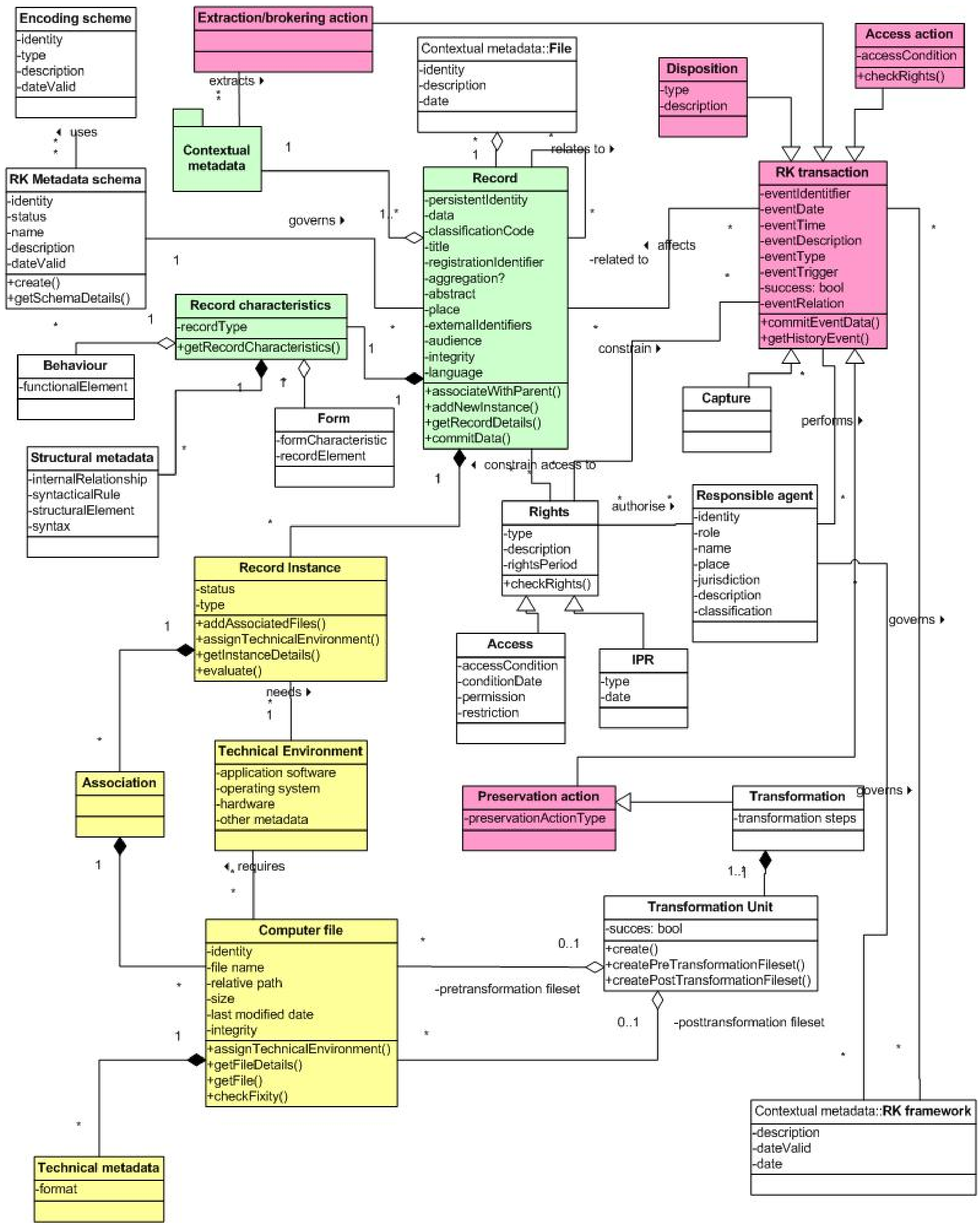


Figure 8. UML Class diagram representing the classes from a recordkeeping perspective

The following main clusters of UML classes can be identified:

1. the intellectual record and its aggregates, the associated contextual metadata and the essential characteristics (blue);
2. the technical components and aspects of the digital record (light yellow); and
3. the recordkeeping transactions (purple).

The broader context: the Ambient Society

In the IDEF0 model, the broader context of the organization is reflected in the A-0 diagram, which identifies how the organization is communicating with both its immediate and more distant environments.

These can be divided into four different areas:

1. developments, laws and other influences that have an impact on how the organization will operate and carry out its mandate. Together, these constitute the controls, which include juridical requirements, corporate culture, recordkeeping theory and methods, mandate, societal influence and state of technology;
2. input (material and immaterial) that are used by the organization and/or transformed into output;
3. output that serves the customers, citizens or other stakeholders; and
4. mechanisms that enable the organization to function, such as staff or personnel, technical infrastructure and facilities.

The top-level IDEF0 diagram, A-0 (see Figure 1), shows the most important influences and relationships between an individual or an organization and its environment. Within the lower level diagrams of the model some of these may be further decomposed into more precise descriptions.

The UML class diagram with the societal view (Figure 8) is an attempt to model a recordkeeping perspective on society. It reflects the fourth dimension of the records continuum paradigm and, as such, is an entity in its own right.[79] It is within society that records, archives and archival institutions play a role and it will also be developments and aspects of society that will determine how records and archives will be created, used and managed.

[79] For a more comprehensive description of the societal aspects of the records continuum, see Sue McKemmish, Michael Piggott, Barbara Reed and Frank Upward (eds.), *Archives: Recordkeeping in Society* (Wagga Wagga, Australia: Centre for Information Studies, 2005), especially chapters 7–12.

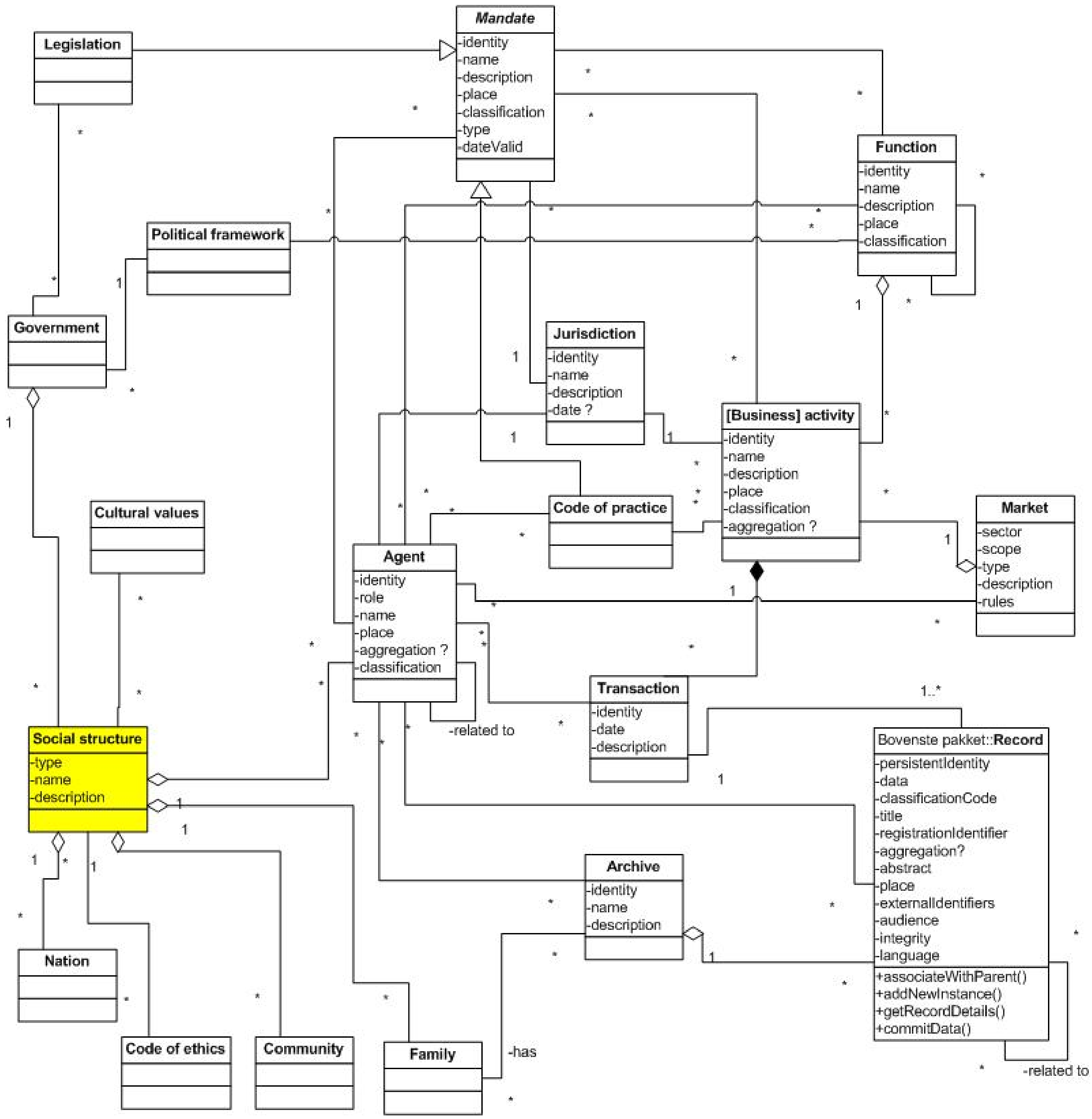


Figure 9. UML Class Diagram representing the societal context

PART SIX

INVESTIGATING THE ROLES AND REQUIREMENTS, MANIFESTATIONS AND MANAGEMENT OF METADATA IN THE CREATION OF RELIABLE AND PRESERVATION OF AUTHENTIC DIGITAL ENTITIES

Description Cross-domain Task Force Report

by
Anne Gilliland, University of California, Los Angeles

Co-authors
Lori Lindberg, University of California, Los Angeles
Victoria McCargar, Los Angeles Times
Alison Langmead, University of California, Los Angeles
Tracey P. Lauriault, Carleton University
Monique Leahey-Sugimoto, University of California, Los Angeles
Joanne Evans, University of California, Los Angeles
Joe Tennis, University of Washington
Holly Wang, University of California, Los Angeles

Introduction

Metadata that are associated with either an information system or an information object for the purposes of description, administration, legal requirements, technical functionality, use and usage and preservation, play a critical role in ensuring the creation, management, preservation, discovery, use and re-use of trustworthy materials, including records. Recordkeeping[1] metadata, of which one key type is archival description, play a particularly important role in documenting the reliability and authenticity of records and recordkeeping systems as well as the various contexts (legal-administrative, provenancial, procedural, documentary and technical) within which records are created and kept as they move across space and time. In the digital environment, metadata are also the means by which it is possible to identify how record components—those constituent aspects of a digital record that might be managed, stored and used separately by the creator or the preserver—can be reassembled to generate an authentic copy of a record or reformulated per a user's request as a customized output package. Metadata in the sciences also provide essential data quality elements such as accuracy, lineage, reliability, margins of error, limitations and precision, among others, that assist the user to assess whether the dataset in questions is fit for the intended use of the scientist.

Issues relating to the creation, capture, management and preservation of adequate metadata are, therefore, integral to any research study addressing the reliability and authenticity of digital entities created by dynamic, interactive and experiential systems, regardless of the community, sector or institution within which they are being created. The Description Cross-domain Task Force examined the conceptualization, definitions, roles and current functionality of metadata and archival description in terms of requirements generated by InterPARES 1 as well as case study data and models generated during InterPARES 2. Because of the needs to communicate the work of InterPARES in a meaningful way across not only other disciplines, but also different archival traditions; to interface with, evaluate and inform existing standards, practices and other research projects; and to ensure interoperability across the three focus areas of InterPARES 2, the Description Cross-domain also addressed its research goals with reference to wider thinking about and developments in recordkeeping and metadata.

InterPARES 2 addressed not only records but also a range of digital information objects (often referred to as "entities" by InterPARES 2, but not to be confused with the term "entities" as used in metadata and database applications) that are the products and by-products of artistic, scientific and governmental activities that are carried out using dynamic, interactive or experiential digital systems. The nature of these entities was determined through a diplomatic analysis undertaken as part of extensive case studies of digital systems that were conducted by the InterPARES 2 Focus Task Forces. This diplomatic analysis established whether the entities identified during the case studies were records, non-records that nevertheless raised important concerns relating to reliability and authenticity or "potential records." To be determined to be records, the entities had to meet the criteria outlined by archival theory—they had to have a fixed documentary format and stable content. It was not sufficient that they were considered to be or were treated as records by the creator. "Potential records" is a new construct that indicates that a

[1] "Recordkeeping" is used in the archival literature in the context of the records continuum to signify an archival worldview of the integration and continual interactivity of processes and responsibilities related both to records creation and to archival management of those records. However, this is not a universally accepted premise, with the lifecycle model drawing a much clearer demarcation between the management of active records and the preservation of archival records. In the Chain of Preservation activity model developed by InterPARES 2, which is based upon the lifecycle model, "recordkeeping" refers to the phase in the lifecycle that comes between "record creation" and "record preservation."

digital system has the potential to create records upon demand, but does not actually fix and set aside records in the normal course of business. The work of the Description Cross-domain, therefore, addresses the metadata needs for all three categories of entities.

Finally, since “metadata” as a term is used today so ubiquitously and in so many different ways by different communities that it is in peril of losing any specificity, part of the work of the Description Cross-domain sought to name and type categories of metadata. The Description Cross-domain also addressed two areas of increasing importance in the digital environment: incentives for creators to generate appropriate metadata; and management issues associated with the retention, maintenance and eventual disposition of the metadata that aggregate in exponentially increasing amounts around digital entities over time.

Research team

The following is a list of researchers and research assistants who participated in the Description Cross-domain Task Force at some point throughout the Project.

Chairs and Co-chairs:

Terry Eastwood	2005-2006 (Chair)
Anne Gilliland	2001-2005 (Co-chair)
Sue McKemmish	2001-2005 (Co-chair)

Researchers:

Martine Cardin	Lavalle University, Quebec, Canada
Terry Eastwood	The University of British Columbia, Canada
Anne Gilliland	University of California, Los Angeles, USA
Hans Hofman	National Archives of the Netherlands
Richard Marciano	San Diego Supercomputer Center, USA
Victoria McCargar	consultant, Los Angeles Times, USA
Sue McKemmish	Monash University, Melbourne, Australia
Joe Tennis	University of Washington, USA
James Turner	Université de Montréal, Canada
Stefano Vitali	Archivio di Stato di Firenze, Italy

Research Assistants:

Bart Ballaux	The University of British Columbia, Canada
Lauren Cardinal	University at Albany, State University of New York, USA
Chia-Ning Chiang	The University of British Columbia, Canada
Joanne Evans	Monash University, Melbourne, Australia
Michael Garabedian	University of California, Los Angeles, USA
David Gibbs	University of California, Los Angeles, USA
John Juricek	University of California, Los Angeles, USA
Eleanor Kleiber	The University of British Columbia, Canada
Alison Langmead	University of California, Los Angeles, USA
Tracey P. Lauriault	Carleton University, Ottawa, Canada
Monique Leahey-Sugimoto	University of California, Los Angeles, USA
Lori Lindberg	San Jose State University and UCLA, USA
Jennifer Osorio	University of California, Los Angeles, USA

Catherine Miller	The University of British Columbia, Canada
Rachel Mills	The University of British Columbia, Canada
Shaunna Moore	The University of British Columbia, Canada
Randy Preston	The University of British Columbia, Canada
Nadav Rouche	University of California, Los Angeles, USA
Wendy Sokolon	The University of British Columbia, Canada
Emily Staresina	University of California, Los Angeles, USA
Stuart Sugarbread	University of California, Los Angeles, USA
Shannon Supple	University of California, Berkeley, USA
Melissa Taitano	University of California, Los Angeles, USA
Holly Wang	University of California, Los Angeles, USA
Eun Young	University of California, Los Angeles, USA
Yuchai Zhou	Carleton University, Ottawa, Canada

Research Questions

Metadata investigations in the digital environment tend to cover a lot of territory, and the scope of the Description Cross-domain as determined in the research proposals funded by the various agencies that supported this work reflect that. The overall work was directed by the questions posed in the Project funded by the Social Sciences and Humanities Research Council (SSHRC) of Canada:

- What is the role of descriptive schemas and instruments[2] in records creation, control, maintenance, appraisal, preservation and use in traditional recordkeeping systems in the three focus areas?
- What is the role of descriptive schemas and instruments in records creation, control, maintenance, appraisal, preservation and use in emerging recordkeeping systems in digital and Web-based environments in the three focus areas? Do new tools need to be developed and, if so, what should they be? If not, should present instruments be broadened, enriched, adapted?
- What is the role of descriptive schemas and instruments in addressing reliability, accuracy and authenticity requirements (including the InterPARES 1 Benchmark and Baseline Authenticity Requirements) concerning the records investigated by InterPARES 2?
- What is the role of descriptive schemas and instruments in archival processes concerned with the long-term preservation of the records in question?
- Do current interoperable frameworks support the interoperability of descriptive schema and instruments across the three focus areas? If not, what kinds of frameworks are needed?
- What are the implications of the answers to the above questions for traditional archival descriptive standards, systems and strategies? Will they need to be modified to enable archival programs to meet new requirements, or will new ones need to be developed? If so, what should they be?
- To what extent do existing descriptive schemas and instruments used in the sectors concerned with the focus areas addressed by this project (for example, the geospatial data community) support and inform requirements such as those developed by InterPARES 1?

[2] This phrase is used throughout to refer to metadata in the broadest sense, as well as to archival description specifically.

Will they need to be modified to enable these sectors to meet these requirements, or will new ones need to be developed? If so, what should they be?

What is the relationship between the role of descriptive schemas and instruments needed by the creator and those required by the preserver to support the archival processes of appraisal, preservation and dissemination? What tools are needed to support the export/import/exchange of descriptive data between systems?

- What is the role of descriptive schemas and instruments in rights management and in identifying and tracking records components, versions, expressions, performances and other manifestations and derivative works?
- Is it important to be able to relate the record of artistic and scientific activity to the associated expression, performance, product, work or other manifestation of it and, if so, in what ways can descriptive activities facilitate it?

Additional research direction came from the projects funded by the United States National Science Foundation (NSF) and the National Historical Publications and Records Commission (NHPRC) that supported the U.S. Team's participation in InterPARES 2. This included formulation and testing of metadata models; and identification of new and existing methodologies and strategies for ensuring that records created using interactive, experiential and dynamic systems can be trusted as to their content (that is, are reliable and accurate) and as records (that is, are authentic) while used by the creator; new and existing methodologies and strategies for selecting records that have to be kept for legal, administrative, social or cultural reasons after they are no longer needed by the creator; new and existing methodologies and strategies for preserving them in authentic form over the long term; and advanced technologies for the implementation of these methodologies in different sectors and disciplinary and socio-cultural contexts. The research was also to develop hypotheses of metadata necessary for prototype systems; and rules for the ongoing description of digital records.

In the course of its work, the Description Cross-domain surfaced and addressed several additional provocative questions:

- Can a vocabulary be created to assist in the identification of different types and functions of metadata?
- What kind of management regime needs to be put in place to ensure the creation and maintenance of trustworthy metadata?
- Can metadata associated with the creation and active use of records ever contribute to archival description, particularly in the capture and elucidation of certain kinds of context and fundamental identification and arrangement information relating to the records?
- Should a metadata specification model generated out of InterPARES 2 support a single or multiple worldviews on the activities, roles, responsibilities and points of engagement with the record (e.g., lifecycle, records continuum and information continuum perspectives)?
- Can metadata-based automated tools support any new kinds of roles and capabilities for the description and use of preserved digital materials?

The latter questions have particular relevance for specifying how the benchmark and baseline requirements developed in InterPARES 1 and discussed further below, are implemented within recordkeeping and archival processes and systems design, as well as for the conceptualization and labelling of the models being developed.

Research Methodology

Multiple, interdependent activities and associated methods were used to generate products and data that could be triangulated to answer the questions outlined above (the researchers primarily engaged in each activity are indicated in parentheses).

- Collecting, compiling and analyzing data on the types and sources of metadata used in real-life dynamic, interactive and experiential systems as identified through case and general studies in arts, science and government settings that were conducted in other InterPARES 2 groups. Method used: *case studies* (focus group case study researchers, UBC project staff, Gilliland).
- Conducting a special investigation to identify state-of-the-art thinking and practice relating to metadata in news archives. Method used: *survey* (McCargar, Supple).
- Developing a database for analyzing warrant (i.e., the mandate from law, professional best practices, professional literature and other social sources) requiring the creation and continued maintenance of description and other metadata supporting the accuracy, reliability, authenticity and preservation of records and other record-like objects. This warrant will be integrated into public recommendations made by the Description Cross-domain and other InterPARES 2 research units with regard to evaluating, extending or revising existing descriptive and metadata schemas; encouraging the creation of meaningful metadata in the arts, science and government; as well as promoting the Metadata Specification Model in systems design. Method used: *literary warrant analysis* (Researchers: Gilliland, Sugarman, Gibbs, Garabedian).
- Developing and compiling a metadata schema registry that unambiguously describes salient features of relevant extant descriptive and other metadata schemas, element sets, standards and application profiles; and identifies existing cross-walks between them. Methods used: *iterative systems design* (Researchers: Gilliland, McKemmish, Hofman, Marciano, Lindberg, Evans, Rouche, Wang, Leahey-Sugimoto, Langmead, Zhou[3]).
- Developing an analytical framework for assessing the extent to which current metadata sets and implementations meet the requirements of the InterPARES benchmark and baseline requirements and/or the ISO Records Management Metadata Standard requirements (subsequently integrated with the registry to create the Metadata and Archival Description and Analysis System (MADRAS); and identifying how such metadata could be extended or modified to meet better recordkeeping requirements. Methods used: *requirements operationalization, warrant analysis, schema analysis, metadata mapping* (Researchers: Gilliland, McKemmish, Hofman, Marciano, Lindberg, Evans, Rouche, Wang, Leahey-Sugimoto, Langmeade, Youn).
- Developing metadata specifications to accompany the activity models constructed by the Modeling Cross-domain. The specifications identify the type, source and application of metadata implicit or explicit in the models and when, how and by whom it should be created.[4] These specifications can also form the basis for developing automated tools (not to be confused with descriptive instruments) that can be used to assist with the creation,

[3] Yuchai Zhou (2005), "Profiling and Visualizing Metadata for Geo-referenced Multimedia Information in a Geospatial Portal: A Case Study for the Cybercartography and the New Economy Project" (Master's thesis, Department of Geography and Environmental Studies, Carleton University, 2005).

[4] The metadata specification model for the Business-driven Recordkeeping Model developed by the Modeling Cross-domain is still to be developed.

capture, management and preservation of essential metadata for active and preserved records. Method used: *modeling and empirical instantiations* (Researchers: Tennis, Eastwood and Preston).

- Interfacing with other relevant research and development activities such as the development of the ISO 23081 Records Management Metadata Standard, the Monash University-based Clever Recordkeeping Metadata Project[5] and the work of the San Diego Supercomputer Center on the development of metadata tools for the automated creation, harvesting and end-user manipulation of metadata (Hofman, Gilliland, McKemmish, Marciano, Evans and Lindberg).

Figure 1 illustrates the relationships between the constituent components and some of the associated activities of the Description Cross-domain. Numbers 1-3 on the flowchart indicate the primary loci of activity and eventual products.

Metadata and Archival Description Registry and Analysis System (MADRAS)[6]

MADRAS was initially envisioned as a metadata registry that could be used by the Description Cross-domain to identify relevant metadata sets and schemas that it wished to evaluate to generate recommendations in response to its research questions. However, it quickly became clear that if the Description Cross-domain was to operate on the assumption that metadata were essential to the creation of reliable and preservation of authentic records in electronic systems of any type, then it also needed to address issues associated with how trustworthy metadata are created and maintained. It was also clear that the Description Cross-domain needed to operationalize the benchmark and baseline requirements generated by InterPARES 1 in terms of how they might be met through metadata and archival description. MADRAS evolved, therefore, beyond being a schema-level (i.e., not a comprehensive element-level) metadata registry, to include an analytical assessment tool that could be used by the researchers to evaluate the current capabilities of registered metadata schemas. With an extension of U.S. research funds until June 2007, it is now envisaged that the beta production version completed in InterPARES 2 and used by the Project's researchers to answer their research questions, will be revised as a full-fledged, publicly available metadata assessment and tracking tool with more sophisticated public interfaces, report formats and privacy controls that will support those who wish to register proprietary or draft schemas.

---

5 See Records Continuum Research Group (1998), "Create Once, Use Many Times - The Clever Use of Metadata in eGovernment and eBusiness Processes in Networked Environments." Available at http://www.sims.monash.edu.au/research/rcrg/research/crm/.

6 For further details on the development of MADRAS, see Anne J. Gilliland, Nadav Rouche, Joanne Evans and Lori Lindberg (2005), "Towards a Twenty-First Century Metadata Infrastructure Supporting the Creation, Preservation and Use of Trustworthy Records: Developing the InterPARES 2 Metadata Schema Registry," *Archival Science* 4(1): 43–78; Joanne Evans and Nadav Rouche (2004), "Utilizing Systems Development Methods in Archival Systems Research: Building a Metadata Schema Registry," *Archival Science* 4(3-4): 315–334; Joanne Evans and Lori Lindberg, "Describing and Analyzing the Recordkeeping Capabilities of Metadata Sets," in *DC-2004: Proceedings of the International Conference on Dublin Core and Metadata Applications, October 11-14 2004, Shanghai, China* (Shanghai, China: Shanghai Scientific and Technological Literature Publishing House, 2004), 75–80. Online reprint available at http://www.dublincore.go.kr/dcpapers/pdf/2004/Paper_27.pdf; Anne J. Gilliland-Swetland and Sue McKemmish, "A Metadata Schema Registry for the Registration and Analysis of Recordkeeping and Preservation Metadata," in *Proceedings of the Second IS&T Archiving Conference, April 26-29, 2005, Washington, D.C.* (Springfield, VA: Society for Imaging Science and Technology, 2005), 109–112; and Lori Lindberg, Monique Leahey-Sugimoto, Nadav Rouche and Holly Wang, "MADRAS: A Metadata and Archival Description Registration and Analysis System for the Analysis of the Recordkeeping Capabilities of Metadata Sets," in *Proceedings of the Third IS&T Archiving Conference* (Springfield, VA: Society for Imaging Science and Technology, 2006), 216–218.

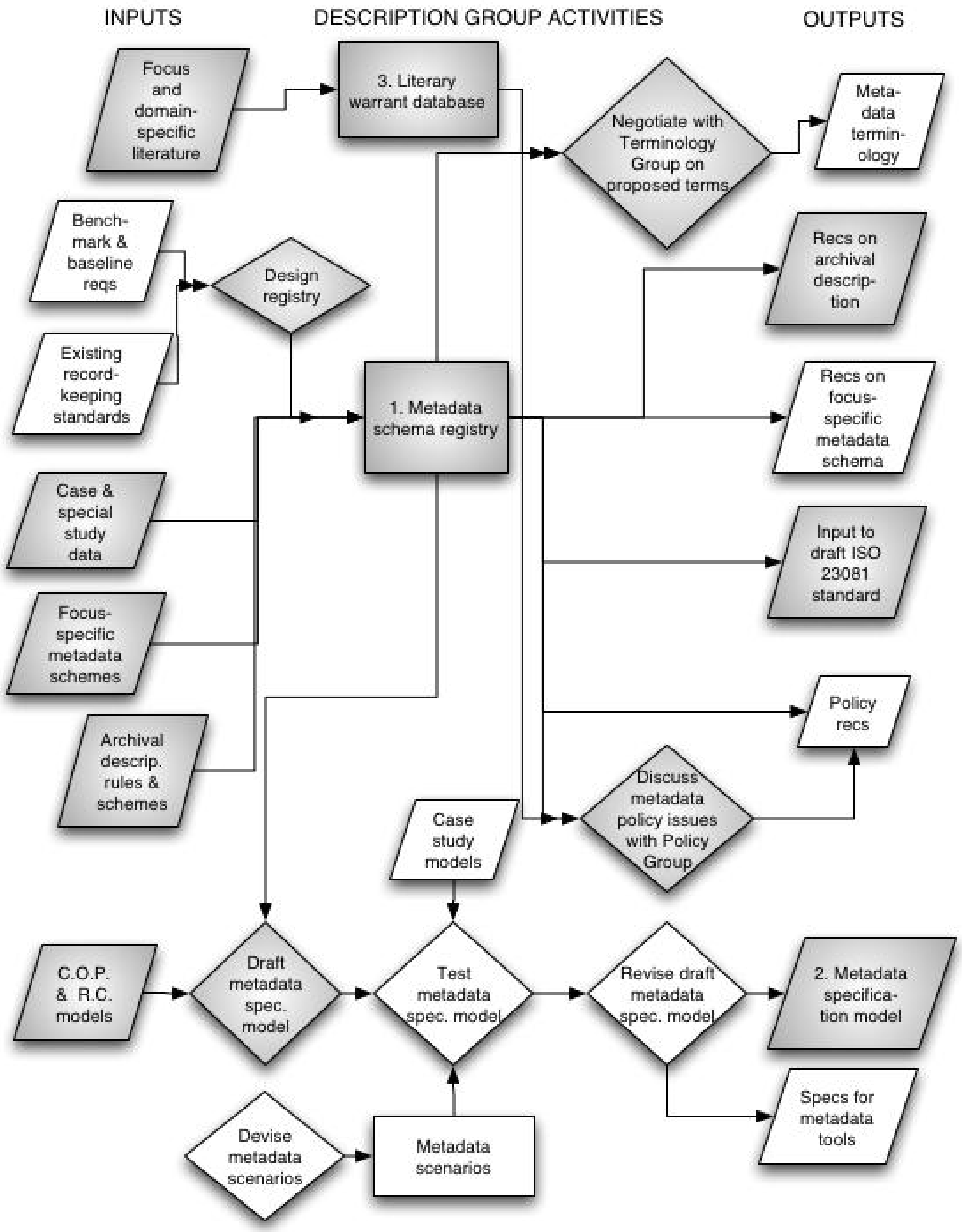


Figure 1. Flowchart of Description Cross-domain Activities

The purpose of MADRAS is fourfold:

1. To support the unambiguous registration of relevant metadata schemas, sets and application profiles;
2. To support the analysis of registered items against requirements derived from the InterPARES 1 benchmark and baseline requirements as well as from the ISO 23081 Records Management Metadata Standard and to make recommendations for how they might be extended or otherwise revised to address the reliability, authenticity and preservation needs of records created within the domain, community or sector to which they pertain.
3. To provide a standardized framework by which any existing or draft metadata schema or set can be assessed for its ability to address the above mentioned requirements and which could be adopted by standards-setting bodies in different areas of practice.
4. To generate analytical data to be provided to the working group (ISO TC46/SC11-WG1) that oversees the development of ISO 23081 for possible incorporation into Part III of that standard.

The inputs for MADRAS development included the following:

a. The benchmark and baseline requirements generated by InterPARES 1.
b. Requirements derived from an analysis of ISO 23081.
c. Requirements derived from analysis of other salient electronic records standards and projects, including the conceptual and relationship models of records in business and socio-legal contexts developed by the SPIRT Recordkeeping Metadata Project and Kate Cumming's "Derivation of the Classification of Recordkeeping Metadata by Purpose Scheme."[7]
d. Metadata schemas and sets identified in the course of the case and general studies undertaken by the focus groups.
e. Other relevant focus-specific metadata schemas and sets identified by the focuses or by the Description Cross-domain (e.g., Geomatics Metadata Standard, ISO 19115).
f. Archival description rules, sets and related practices (e.g., ISAD(G)/ISAAR, EAD/EAC/ DACS, RAD and the Australian Series System).

The current beta environment for MADRAS is implemented using PHP, a server-side scripting language that provides Web development tools for building dynamic Web sites. The back-end Web server is Apache 1.3 and the database server is MySQL 3.22. Both servers are hosted on a machine running the Unix operating system. PHP, Apache and MySQL are all open-source technologies and are used by many database-driven Web applications. Information about the process of building MADRAS has been kept in MADRAS itself using an online note sharing system. The current size of MADRAS is 20 megabytes (without appended documents) with around 100 PHP files. More files will be generated in conjunction with the development of the analysis interface. The researchers expect that MADRAS will grow into a mid-sized application after processing more feedback from InterPARES researchers and adding more data and infrastructure. MADRAS is allowed 50,000 queries per hour from the database server and MySQL 3.22 has a 4-gigabyte limit on table size (limitations are a function of MySQL).

[7] See Kate Cumming, "Purposeful Data: The Roles and Purposes of Recordkeeping Metadata" (Ph.D. dissertation, Monash University, 2005).

MADRAS registry component and the schema registration framework

As Chris Hurley has noted:

> Contextual metadata documents circumstances relevant to the making of the record: who, when, how, why … Efforts now being made to regularize the process whereby knowledge of context is captured as metadata for electronic record-keeping should not blind us to a fundamental truth. Because records themselves are timebound, metadata must be verified within a context which is both current *and* historical. Records cannot remain current unless the metadata is externally validated.[8]

Hurley is arguing that beyond the comprehensive and rigorously delineated metadata and archival description necessary for creating reliable records and maintaining and demonstrating the authenticity of archival records, there is a need for overt integrity control and transparency of those metadata and of archival description. This can only be the case if the metadata themselves are trustworthy and comprehensively managed for as long as they are required. In other words, reliability and authenticity are concerns for recordkeeping metadata as well as for the records and recordkeeping processes to which they relate. Metadata generated and managed by records creators and archival description generated by archivists, must be sufficient, appropriate, understandable and of high quality. MADRAS and the metadata specification model, therefore, are two tools that seek to support a highly reflexive recordkeeping metadata regime that addresses both of these concerns.[9]

The MADRAS registry component was developed with the following primary purposes in mind:

- to describe relevant metadata schemas and their features in a standardized way;
- to provide an overview of existing and emerging schemas;
- to provide an overview of the applicability of the schemas to recordkeeping and archival functions;
- to describe the scope and purpose of the schemas;
- to specify what type(s) of metadata they cover; and
- to identify related schemas (e.g., schemas that control data values, schemas that provide structure for metadata elements).

With one of the expected outcomes of the Description Cross-domain research within InterPARES being a production of “scholarly comparative discussions of existing descriptive standards and an intellectual framework of descriptive standards for the records under examination [within InterPARES 2],”[10] MADRAS was developed to act as a data collection and analysis tool for InterPARES 2 researchers. After developing a framework for the standardized description of metadata schemas, a metadata schema itself was produced in the form of an XML DTD. From the DTD, a prototype database was developed to assist researchers in the refinement of the DTD

[8] Chris Hurley (1995), “Ambient Functions: Abandoned Children to Zoos,” *Archivaria* 40 (Fall): 22. Emphasis in original. Online reprint available at http://journals.sfu.ca/archivar/index.php/archivaria/article/viewFile/12095/13080.

[9] Archives have always been metadata-rich environments, although they are not always recognized as such, just as archival description is not always recognized by archivists as the primary means by which they demonstrate the authenticity of their holdings. Archivists must be cognizant that the accession records, finding aids and use records they typically create today are not only part of the archival description for the records to which they relate, but they are also records in their own rights. The scrutiny, therefore, that archivists give to the records and recordkeeping metadata of others to assess and validate their management and reliability, they must also give to their own.

[10] InterPARES 2 Project, “Overview of InterPARES 2 Intellectual Framework,” 7. Available at http://www.interpares.org/display_file.cfm?doc=ip2_overview_of_intellectual_framework(20030311).pdf.

design. Now that the production version of the registry is operational, the prototype database has been retired. MADRAS is the result of the lessons learned from the prototype database.

The following outlines the development process and design decisions involved in the building of MADRAS:

- The decision to develop the registry as a way to approach the Description Cross-domain research was based upon the realization that it was impossible to assess all relevant schemas within the time available to the Project and also that any such assessment would date rapidly, given the current pace of schema evolution. Instead, researchers decided to develop a tool that could be used into the future by any party wishing to register and assess a schema they were using or planned to use against InterPARES requirements, as well as discover other schemas, view them and their assessments and learn about them in a consistent, structured environment. This decision is significant because it reflects a pragmatic approach to the political realities of metadata schema creation and use. Schemas have proliferated in many communities and are closely tailored to their specific needs. The Description Cross-domain decided that it was very unlikely that any community would adopt a schema developed by InterPARES in place of or in addition to its own. Instead, the approach adopted demonstrates how interested parties can use their own community or implementation-specific schemas, compare them to others and begin to think about them in the larger world of metadata schema development as well as in different ways. This reflexive thinking was evident in the Description Cross-domain researchers' own discovery processes when they decided additionally to address in the MADRAS analysis component, the requirements contained in the ISO 23081 standard, so that users could both assess their schemas and compare differences between recordkeeping metadata requirements as articulated by the InterPARES Project research and another research collaboration, that within the ISO.
- The first step toward developing a registry was to develop a draft XML Document Type Definition (DTD) that would become the backbone of the registry. XML was chosen because of its platform independence, flexibility at handling hierarchical data and relative ease of migration. In its original conception, MADRAS was to be an integrated description and analysis tool, with all data encoded within the DTD. The researchers decided to move ahead with the analysis component of MADRAS simultaneously using an Excel spreadsheet-based worksheet, allowing for parallel work activities while the DTD was being tested and a prototype registry built. A form of the analysis worksheet was originally intended for integration into the DTD. However, the analysis became such a large and significantly complex component it was not integrated into the DTD and became an independent part of the registry.
- To develop the DTD, the researchers examined how metadata about metadata schemas should be sourced to ensure their reliability and authenticity, for example, through recordkeeping requirements for metadata registries described in the ISO standard for metadata registries.[11] In addition, the researchers were mindful that the description data within the registry component was not required to conduct an in-depth analysis of a schema but rather to extract structured objective information about a schema as it is described in schema documentation; for example, an official name of a schema, its acronym, publisher information, documentation pointers or citation information and copyright statements.

[11] International Organization for Standardization, International Electrotechnical Commission, ISO/IEC 11179: Information Technology—Metadata registries (MDR). Available at http://metadata-stds.org/11179/index.html.

- The registry DTD was developed with a classification hierarchy of elements three levels deep. Level one elements corresponded to major sections of metadata about metadata schemas (hereafter descriptors):
  - REGISTRATION
  - IDENTIFICATION
  - TECHNICAL REQUIREMENTS
  - RIGHTS
  - PROVENANCE
  - DESCRIPTION
  - DOCUMENTATION
  - RELATIONSHIPS
  - NOTE

  Each of the Level one elements (except NOTE) possessed one or more Level two sub-elements and some Level two sub-elements possessed one or more Level three sub-elements.
- Once the registry DTD was developed, the researchers identified multiple key metadata and descriptive schemas and sets (both from the archival field and from those in use in sectors within the three InterPARES 2 focus areas—arts, science and government) and registered them in a prototype database to test and refine MADRAS database. This prototype database, built in Microsoft Access, allowed the Description Cross-domain researchers to view the registry records and test the DTD as an encoding standard for the registration and description of metadata schemas. Visualizing and working with the data input as well as with the descriptors (the DTD elements) in the prototype registry allowed for the rearrangement of DTD elements, identification of mandatory elements and the proposal of possible controlled vocabularies for certain element values, along with the identification of the need for particular element data value encoding schemas such as ISO 8601 for representation of dates and other relevant ISO standards for forms of country names, languages and so on.
- Guidelines for registering and describing schemas were developed and refined as the DTD researchers' experience with the system increased. To test these guidelines and to check for intercoder consistency, graduate students who had not previously been involved in MADRAS development were assigned schemas to register. In addition, students from other research laboratories not familiar with archiving metadata methods also assisted with the registration of their discipline-specific metadata standards. This provided valuable feedback from persons who were viewing the registry for the first time and who were not necessarily from recordkeeping backgrounds.
- Documentation of system functionality and requirements was developed to support the ability to maintain the system and facilitate the eventual transfer of it from UCLA, where it was developed, to a maintenance agency.

The registration process for MADRAS involves manual entry of values into templates. Population of the prototype database demonstrated wide variation in how schemas are published and information about them is presented. In such circumstances, manual processes involving human cognition, collation and data entry appear to be the only viable registration method, since humans are best able to negotiate the situation-specific mappings and cope with gaps and ambiguities in the schemas. Utilizing such an approach, however, also introduces scalability and sustainability issues for MADRAS, given the amount of manual processing required. It points to the need for standardization in the way metadata standards, schemas, crosswalks and their meta-information are published so that registration can be automated or at least semi-automated. This also raises the question of what meta-information should be made available as part of the publication of metadata standards and for the consumption of what types of agents.

The development of the registry component of MADRAS was an iterative process that was striving for an ideal system. Even through the tremendous intellectual capital invested in the prototype database, the DTD was not able to address fully all of the researchers' questions, and the difficulties in so doing provided valuable insight into some of the metadata issues being addressed. Additional valuable questions were raised during the transformation of the DTD and the prototype into the production version of MADRAS and the building of the technological infrastructure for distributed registration and analysis activity in the Web environment.

Analytical framework

The analytical component of MADRAS was developed through iterative prototyping and warrant analysis over a period of three years. The technique of warrant analysis was employed to determine the criteria against which judgments as to the recordkeeping and archival capabilities of metadata schemas could be made. The process involved studying each warrant for statements made regarding requirements for recordkeeping metadata and turning these into a series of questions. These questions were then compiled into an analysis worksheet using an Excel spreadsheet. Although there was a degree of overlap in these statements, the strategy was to have separate sections for each warrant as part of the data gathering that would feed into the metadata model developments.

A primary set of conditions against which metadata schemas registered in MADRAS are assessed is the benchmark and baseline requirements that were generated out of the InterPARES 1 Project.[12] The benchmark requirements are based on the notion of a trusted recordkeeping system. They include requirements that support the presumption of the authenticity of digital records before they are transferred to the preserver's custody. The baseline requirements are based on the notion of the preserver as trusted custodian and support the production of authentic copies of digital records after they have been transferred to the preserver's custody. These are the only extant sets of requirements that specifically address how creators and archivists can assess the authenticity of records. As noted by Evans and Lindberg:

> The benchmark requirements identify the record attributes (metadata) that need to be 'explicitly expressed and inextricably linked' to a record in order for its identity and integrity to be asserted. The benchmark requirements also identify 'the kinds of procedural controls over the record's creation, handling and maintenance that support a presumption of its integrity.' The role of the benchmark requirements is to act as a tool for preservers to use in assessing the authenticity of electronic records. The higher the number, and the greater the degree to which a system meets these requirements, then the stronger the presumption of the authenticity of the electronic records held within it.

[12] See Authenticity Task Force, "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records" in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 204–219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf. Abridged versions of the benchmark and baseline requirements are provided in Appendices 21a and 21b, respectively. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_21.pdf.

> In contrast, the baseline requirements specify the requirements that must be met in order to produce authentic copies of electronic records from a preservation system. This includes archival descriptive metadata documenting 'the records juridical-administrative, provenancial, procedural and documentary contexts,' and controls over the records transfer and reproduction processes to ensure the maintenance of the records' identity and integrity.[13]

As this excerpt indicates, many of the benchmark requirements could potentially be implemented through metadata and archival description, particularly such aspects as identity, linkages, documentation of documentary forms, juridical requirements, business rules and technical procedures, access privileges, establishment of the authoritative record when multiple copies exist and transfer of relevant documentation; as could almost every aspect of the baseline requirements. The benchmark and baseline requirements, however, had only been expressed conceptually, and in narrative form, by InterPARES 1 and were not operationalized for any kind of technological implementation; for example, as a set of logical propositions or production rules. Nor were the requirements deconstructed in a way that would specify how other processes and metadata might help to meet them. For example, how might the different types of context identified in InterPARES 1 be manifested or documented through metadata? One way of addressing a problem such as this is to decompose archival and recordkeeping notions of "context" into types that can then be associated with specific processes and attributes. InterPARES 1 identified five different types of contexts as being relevant to the maintenance of authentic records over time: juridical-administrative, provenancial, procedural, documentary and technological.[14] Some of these types need to be further decomposed to identify their constituent metadata manifestations.[15]

Accordingly, the development of the analytical framework used in MADRAS sought to operationalize these narrative requirements in terms of how they might be satisfied both through the metadata associated with the active record and recordkeeping system and archival description. The same then had to be done for the ISO 23081 requirements, which were also narratively expressed. Once the framework was drafted, Description Cross-domain researchers analyzed multiple existing schemas, standards and guidelines to assess the extent to which they met the requirements, given their stated scope. Where the analysis indicates that a schema falls short, the output report generated by MADRAS delineates exactly where and how and researchers can then recommend augmentations or modifications to ensure that the schema meets those requirements that fall within its stated scope. MADRAS can also be used to identify potential companion metadata schemas that can be used to address those parts of the requirements that are unaddressed because they are out of scope (e.g., because the schema addresses the creator or the preserver side only, or is content rather than context- or recordkeeping-centric). When the beta system becomes publicly available, anyone will be able to

[13] Evans and Lindberg, "Describing and Analyzing the Recordkeeping Capabilities of Metadata Sets," op. cit.

[14] See Authenticity Task Force (2001), "Appendix 1: Template for Analysis," in Duranti, *Long-term Preservation*, op. cit., 198–203. Online reprint available at http://www.interpares.org/book/interpares_book_j_app01.pdf.

[15] For example, the juridical-administrative type could potentially be decomposed to address specific types of juridical-administrative requirements that manifest themselves directly in emerging metadata initiatives, such as those relating to rights management for records. Digital rights management (DRM) metadata are increasingly being integrated into systems by creators, publishers and information providers, for example, as mechanisms for expressing and automatically enforcing rights and licensing requirements relating to information resources. In an age where records are more and more often the product of private activity or collaboration or of outsourcing relationships between government and the private sector or academic research, collaborative science and industry, such developments not only reflect these changes in records creation but can have significant implications for both researchers and the types of preservation regimes to which the records may be subject.

register and evaluate a current or draft schema or application profile. In this way, the analytical framework can be applied beyond the duration of the InterPARES 2 Project to assess schemas, sets and application profiles as they develop and evolve. This approach also ensures that multiple models for managing records can be supported—both those that seek to apply an end-to-end recordkeeping metadata schema and those where different parties have responsibility for different aspects of recordkeeping and archival preservation.

To draw on as many perspectives as possible and to try to identify where there might be consensus or divergence about relevant recordkeeping requirements (especially where there might appear to be differing viewpoints emerging from the lifecycle and records continuum perspectives), several other prominent standards, guidelines and requirements were also consulted, including ISO 15489 Information and Documentation—Records Management (2001), the U.S. Department of Defense's Design Criteria Standard for Electronic Records Management Software Applications (DoD 5015.2-STD, 2002), and the European Union's Model Requirements for the Management of Electronic Records (MoReq) that specifies requirements for Electronic Records Management Systems (ERMS).

The decomposed requirements were conceived and expressed in the analytical framework in the form of evaluative questions, with the questions designed primarily to elicit a positive or negative response. For positive responses, a schema's element or elements that satisfied a particular question could be noted. The original Excel spreadsheet was organized to systematically describe schemas and assess them over seven sections: (1) General; (2) Recordkeeping General; (3) ISO 23081; (4) InterPARES benchmark requirements; (5) InterPARES baseline requirements; (6) Classification of Purpose of Recordkeeping Metadata; and (7) General Comments.

The questions were then coded to specific sections of these two instruments so that an actual analysis could be performed.[16] The structure of the worksheet, the nature of the individual questions and the analysis process as a whole was defined and refined through iteration and testing. The questions were applied to a sample of schemas to determine their feasibility, granularity and usefulness as well as the meaning of the response. Schemas included in the sample were selected on the basis of being able to help in determining whether the analysis could make distinctions between recordkeeping and non-recordkeeping schemas, between "single" and "multi-entity" schemas and between schemas operating in different dimensions.

The first attempt to organize the analysis questions was based on a view of what metadata are supposed to do (for example, describe record content, context and structure and then recordkeeping activities). However, to facilitate user comprehension, it was eventually decided to separate the questions by the different recordkeeping entities suggested by the instruments: Record, Agent, Mandate, Business Process and Recordkeeping. To do this, the researchers employed an iterative development process, focused on refining and arranging questions. They paid careful attention to the ways in which each instrument used its own terminology and brought that forward into the analysis questions.

The initial statement of requirements was progressively refined through the development of a prototype database and its population, with a sample of metadata schemas. This process helped to ensure that a flexible descriptive schema was developed that could cope with the diversity of metadata schema publication and documentation practices. It also enabled the testing of the feasibility and applicability of the proposed elements and determination of the sources of metadata values.

[16] See Appendix 17.

The researchers decided that the first iteration of the system would be for InterPARES' researchers themselves and then the system should be revised for future use by other interested stakeholders (i.e., records keepers, archivists, etc.). The analysis worksheet underwent a number of versions and changes through the initial testing and validating that resulted in a final accounting of four major versions of the worksheet with smaller subversions (4.1, 4.2 and so forth). First, the analysis was mocked-up in Excel. Later, during the design development phase, FileMaker Pro was used to work up a model for the display of information in MADRAS that was eventually recreated in the actual MADRAS system.

Challenges encountered in the development of the analytical framework ranged in complexity. Often, it was necessary to return to first principles. For example, during the process of creating MADRAS, researchers needed to come to an agreement (or not) on the operational meaning of the word "record." What did they consider to be a record? A relationship? Along the same lines, researchers needed to consider what the base unit of analysis should be (in other words, to what level of granularity should the analysis proceed?). In the end, the decision was made that the system would proceed to the element and not to the sub-element level.

The researchers experimented with developing various versions of a decision tree. Lacking consensus, they decided not to use any of the versions in the current production version, but did agree to revisit the use of a decision tree in a later version. The process did, however, help with the decision to push certain questions to the registry and table relationships in the analysis until it was decided whether or not a relationship should be elevated to its own entity. Some of the other activities involved in the framework development included the following:

- mapping between related InterPARES and ISO requirements;
- developing controlled vocabularies for classifying the purpose of schema and standards, and for types of metadata specified in schema and standards (drawing on ISO 23081, the SPIRT Recordkeeping Metadata Research Project outcomes and the Records Continuum and InterPARES Models).; and
- exploring the boundaries between and around records and related metadata, and noting that some metadata relate to the content, structure (documentary form) and business context of the record (concerned with the nature of the business transaction captured in the record), and that some metadata relate to the recordkeeping processes that manage the record.

The analysis worksheet stayed fairly stable until the spring of 2005, when the shift from the manual worksheet-based analysis to an automated version of the analysis began. The automation of the analysis process, a goal of the MADRAS tool development, surfaced a number of procedural and technical considerations, not the least of which was the time spent on manual analysis and the time spent to teach new analysts how to do the work. Research team members observed that the analysis reference instruments had a number of areas of overlap and that as a result similar questions that sought similar answers were asked over more than one section of the spreadsheet. The decision was made to map each of the reference instruments against one another to take advantage of commonalities amongst the instruments. This decreased the amount of repetitive work, as well as verified for the researchers that the research findings across the different projects producing the reference instruments came to some common conclusions. For example, when considering the *Classification of Purpose of Recordkeeping Metadata* schema developed by Kate Cumming, the researchers looked very carefully at her classification schema and where it might be expressed or assumed as the basis for requirements expressed in the remaining analysis reference instruments. Cumming concludes that all recordkeeping metadata are created to satisfy one of seven particular purposes:

1. unique identification;
2. authentication of records;
3. persistence of records content, structure and context: by fixing their content, ensuring that their structure can be re-presented, and maintaining sufficient organizational and functional context to preserve their meaning over time and beyond their context of creation;
4. administering terms and conditions of access and disposal;
5. tracking and documenting use history, including recordkeeping and archiving processes;
6. enabling discovery, retrieval and delivery for authorized users; and
7. restricting unauthorized use.[17]

It was determined that these purposes were all articulated in the warrants in one way or another and did not need explicit consideration as a separate grouping of questions in the analysis. The mapping of the reference instruments decreased the number of questions asked in the analysis, making the process more efficient and less time-consuming. In addition, it allowed the analysts to be able to look at the data produced in new ways and apply findings more broadly.

Automating the process of analysis also required re-thinking how consistency could be ensured across different analysts. The researchers were trying to automate a system that relied on an unknown: the extent of the human analyst's knowledge; and this raised interesting issues. The original method of analysis using Excel spreadsheets had demonstrated that analysis could vary considerably according to the knowledge and experience of the analyst. The researchers had to assume certain pre-existing knowledge on the part of the user of the system. It was decided that users would most likely be experienced records keepers or those familiar with archival terminology.

During the automation process, the strengths and weaknesses of the original analysis spreadsheets were assessed to clarify and bolster the effectiveness of MADRAS. This surfaced several issues with the original spreadsheets including that:

- The original worksheet facilitated documenting, rather than analyzing, a metadata schema. (Solution: focus on analyzing rather than on documenting the schema.)
- The original worksheet was repetitious. Information documented in one section was repeated in another. (Solution: eliminate redundancy.)
- The original worksheet and evaluation instruments had confusing language. (Solution: simplify and add documentation. For example, a definition file was created that strives to provide a single definition of terminology to assure analyst consistency.)
- The original worksheet was in a format that did not transfer easily to a database/online worksheet. (Solution: creation of an environment that was flexible enough to experiment with—a FileMaker prototype was created as a design sandbox.)
- The criteria for ranking schemas and evaluating answers were not clear. (Solution: create a system where as much ambiguity as possible could be eliminated.)
- The original analysis process did not allow for the discovery of other relevant types of metadata that might be present in a schema but not in any of the analysis instruments. (It was not possible to address this as the analysis was so strongly focused on the InterPARES and ISO instruments.)
- Although the original analysis process asked for repeatability and the obligation value for each element, the Excel worksheet did not ask for this information. (Solution: The researchers separated out the repeatability (or lack thereof) of fields as well as whether a field is mandatory into the element registration process.)

[17] See Cumming, "Purposeful Data," op. cit.

As the design process continued, the researchers conducted a series of user tests, which generated quite a bit of feedback used to improve the design of the system. They also focused on the creation of a tool where users answer questions about a schema and indicate precisely what elements the schema uses to fulfil a specific requirement.

The researchers attempted to confront the issue of how one separates what is explicitly stated in schema documentation and what is implicit, since they wished to create a tool that would test for the *explicit* nature of the metadata. This issue arose from the following section of ISO 23081:

> Records management has always involved the management of metadata. However, the digital environment requires a different expression of traditional requirements and different mechanisms for identifying, capturing, attributing and using metadata. In the digital environment, authoritative records are those accompanied by metadata defining their critical characteristics. These characteristics must be explicitly documented rather than being implicit, as in some paper-based processes.[18]

ISO interactions

Hans Hofman, National Archives of the Netherlands, served as both a member of the Description Cross-domain and as a member of TC46 SC11 WG01, the Technical Committee overseeing ISO 23081 development. He provided input to and feedback on the development of the registry and the analytical framework from the ISO perspective. One of the MADRAS developers, Lori Lindberg, also travelled to Paris to present the MADRAS work and get feedback directly from the Technical Committee. The feedback from that presentation was that the framework was too “record-centric,” and so the researchers revised the framework somewhat to be more entity-focused.

MADRAS has been developed and constructed by researchers with varying knowledge of records and recordkeeping and drawn from disparate recordkeeping philosophies. Enduring challenges include how to accommodate the various audiences and communities that may utilize MADRAS and providing a transparency of the analysis process to accommodate those without a recordkeeping background who are concerned about these issues but are relatively unfamiliar with recordkeeping theory, processes and terminology. Another, more significant, challenge is how to construct and present questions that address the complexity of the metadata model behind ISO 23081 and the conceptual entities incorporated within the standard in a user-friendly manner. As the metadata counterpart to ISO 15489, the international records management standard, ISO 23081 is in itself quite detailed and complex, with multiple types of metadata accruing at various layers and at different times within a recordkeeping system. With ISO 23081 incorporating the significant findings about the authenticity of records developed within the InterPARES Project as well as the conceptual recordkeeping model behind the Australian Recordkeeping Metadata Standard, itself the basis for ISO 15489, this assessment tool must accommodate both of the major models of records management currently in use in the archives and records management communities—the lifecycle model as reflected in the InterPARES research and the continuum model developed in Australia.[19]

[18] International Organization for Standardization, ISO 23081-1:2006 - Information and documentation—Records management processes—Metadata for records—Part 1: Principles, 2. It is also worth noting, however, that the question may be less a matter of the latency or explicitness of schema documentation than the shortcomings of using tools, such as plain XML, that do not allow for the specification of semantic constraints of entities and the relationship(s) between entities.
[19] See Lindberg et al., “MADRAS,” op. cit.

Data and data analysis

A list was generated of major metadata schemas and sets that are in use in the archival field as well as in the areas covered by each InterPARES 2 focus areas. These include:

- *ANZLIC Metadata Guidelines: Core Metadata Elements for Geographic Data in Australia and New Zealand*—defines metadata elements that describe characteristics of spatial datasets maintained in the public and private sectors;[20]
- *Arizona Electronic Recordkeeping Systems (ERS) Guidelines - IV Functional Requirements for Recordkeeping Systems*—describes specifications for recordkeeping functionality that should be incorporated into any digital information system to ensure it can produce records that are accepted as evidence, well managed and preserved, and that benefits are appropriate to the costs;
- *Australian Recordkeeping Metadata Schema (RKMS)*—defines a highly structured set of metadata elements that conforms to a data model based on that developed for the Resource Description Framework (RDF) and that is designed to be extensible and can inherit metadata elements from other schemas;[21]
- *CURL Exemplars in Digital Archives (CEDARS) Metadata for Digital Preservation*—defines a metadata specification based on the Open Archival Information System (OAIS) model that is designed to be used within the Cedars demonstrator services and as a contribution to international efforts at standardization on preservation metadata;[22]
- *Digital Rights Expression Languages (DREL)*,[23] *Online Information Exchange (ONIX) Metadata Specification*[24]—an international standard for representing and communicating book industry product information in electronic form;
- *eXtensible rights Markup Language (XrML)*—a general-purpose, XML-based specification grammar for expressing rights and conditions associated with digital content, services or any digital resource;[25]
- *Global Information Locator Service (GILS)*—an open standard for searching basic information descriptions based on the ISO 23950 search standard;[26]
- *ISO 19115:2003 Geographic Information—Metadata (geomatics metadata standard)*—defines the metadata elements required for describing geographic information and services, including the identification, the extent, the quality, the spatial and temporal schema, spatial reference and distribution of digital geographic data;[27]
- *ISO 82045-2 Document Management—Part 2: Metadata elements and information reference model*—provides a comprehensive set of standardized metadata elements for document management, including data exchange and implementation of a document management system;[28]

[20] Available at http://www.anzlic.org.au/download.html?oid=2358011755.
[21] Available at http://www.sims.monash.edu.au/research/rcrg/research/spirt/deliver/index.html.
[22] Available at http://www.leeds.ac.uk/cedars/MD-STR~5.pdf.
[23] Available at http://www.jisc.ac.uk/uploaded_documents/TSW0603.pdf.
[24] Available at http://www.editeur.org/onix.html.
[25] Available at http://www.xrml.org/.
[26] Available at http://www.gils.net/index.html.
[27] Available at http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=26020.
[28] Available at http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=34513.

- *MAchine-Readable Cataloging (MARC)*—defines a data format by which computers exchange, use and interpret bibliographic information;[29]
- *Metadata Encoding and Transmission Standard (METS)*—a standard for encoding descriptive, administrative and structural metadata regarding objects within a digital library, expressed using XML;[30]
- *Minnesota Recordkeeping Metadata Standard*—defines metadata elements developed to facilitate records management by government entities at any level of government;[31]
- *New South Wales Recordkeeping Metadata Standard (NRKMS)*—describes metadata that can be used by NSW public sector bodies to meet the business, accountability and archival requirements for records; based on the principles of AS 4390: 1996 (the Australian standard for records management);[32]
- *NISO Z39.87-2002 AIM 20-2002 Data Dictionary—Technical Metadata for Still Images*,[33] *Metadata for Images in XML (NISO MIX) Schema*[34]—defines a set of metadata elements for raster digital images to enable users to develop, exchange and interpret digital image files;
- *NLA Pandora AGLS Metadata Element Set*—metadata elements designed to improve the visibility, accessibility and interoperability of online information and services;[35]
- *Open Digital Rights Language (ODRL)*—a proposed language for the Digital Rights Management (DRM) community for the standardization of expressing rights information over content;[36]
- *PREMIS Data Dictionary for Preservation Data*—defines and describes an implementable set of core preservation metadata with broad applicability to digital preservation repositories;[37]
- *Preservation Metadata - Networked European Deposit Library (NEDLIB) Metadata for Long Term Preservation*—defines preservation metadata elements for a deposit system for electronic publications largely based on the OAIS model;[38]
- *Preservation of Electronic Records in a Records Management Application (PERM) Preservation Attributes*—designed for managing a persistent archives of electronic records created by desktop applications through use of an XML Archiving & Packaging Tool (XAPT);[39]
- *Record Keeping Metadata Requirements for the Government of Canada*—defines metadata elements that identify the type of information Departments are required to capture to describe the identity, authenticity, content, context, structure and management requirements of records created in the context of a business activity;[40]

---

[29] Available at http://www.loc.gov/marc/.
[30] Available at http://www.loc.gov/standards/mets/.
[31] Available at http://www.mnhs.org/preserve/records/metadatastandard.html.
[32] Available at http://www.records.nsw.gov.au/recordkeeping/nsw_recordkeeping_metadata_standard_4614.asp.
[33] Available at http://www.niso.org/standards/standard_detail.cfm?std_id=731.
[34] Available at http://www.loc.gov/standards/mix/.
[35] Available at http://ww.naa.gov.au/recordkeeping/gov_online/agls/metadata_element_set.html.
[36] Available at http://odrl.net/1.1/ODRL-11.pdf.
[37] Available at http://www.oclc.org/research/projects/pmwg/premis-final.pdf. Several InterPARES 2 researchers, in particular Victoria McCargar, were involved with the development of the PREMIS metadata set, which occurred concurrent with the work of InterPARES 2.
[38] Available at http://nedlib.kb.nl/results/NEDLIBmetadata.pdf.
[39] Available at http://www.npaci.edu/online/v6.2/perm.html.
[40] Available at http://www.imforumgi.gc.ca/documents/2001/meta/meta00_e.asp.

- *Recordkeeping Metadata Standard for Commonwealth Agencies*—describes the metadata that the National Archives of Australia recommends should be captured in the recordkeeping systems used by Commonwealth government agencies;[41]
- *South Australian Recordkeeping Metadata Standard (SARKMS)*—a technical standard outlining the basic core set of metadata elements required to capture and maintain recordkeeping metadata to assist in meeting the requirements of providing adequate records management;[42] and
- *Victorian Electronic Records Strategy (VERS) Metadata Schema*—defines metadata elements for a system that will capture, archive, manage and provide access to reliable and authentic electronic records.[43]

Schemas were initially selected based on the following processes and criteria:

- Schema documentation was reviewed and checked for relevance to recordkeeping and/or to see if it would be appropriate to analyze.
- Schemas that did not have sufficient documentation were removed from the list.
- Any schema that was listed as a "crosswalk" was removed since the system was not designed to analyze crosswalks (although the existence of crosswalks related to schemas was noted in the analysis).
- Because of time constraints, schemas that had a very large number of elements were given a lower priority.

The researchers decided that it would be too time-consuming to enter all of the elements of an individual schema manually. For some schemas (such as VERS) that have a large number of schemas *and* that have elements categorized according to a schema, the researchers decided that to enter just the name of the element container and specify the element that satisfied the condition in a note field. In future, an "import" function might be added to collect this data automatically from electronic versions of the schemas instead of having to do it all manually.

From this initial selection, schemas were prioritized based on their type. Schemas typed as those intended for recordkeeping purposes were given high priority. These generally were schemas for either local governments (for example, Minnesota) or for national recordkeeping purposes (for example, the Australian RKMS). Since there were such a high number of schemas for government recordkeeping, the researchers also tried to prioritize by sector. Schemas relating to InterPARES focus areas, such as the arts or geospatial applications, were given a higher priority than others.

One thing the researchers noticed during the prioritization process was that *all* of the schemas were from English-speaking countries, apart from one that was developed in China. It would be interesting to try to find more schemas developed by non-Anglo communities and try to analyze those. The researchers also noted that among the selected schemas, there was not a wide variety by domain/sector registered in the system. Schemas for the legal or medical fields were not represented, for example. It would be useful to get a sampling of these schemas for comparison, especially to see if other relevant metadata were revealed.

Other considerations included weighing what might be gained from analyzing schemas that were not developed specifically for recordkeeping purposes. How do they differ? Are the schemas that were not designed for recordkeeping purposes all necessarily deficient when assessed against recordkeeping requirements? Again, might they include elements not previously considered that might be useful for recordkeeping purposes?

[41] Available at http://www.naa.gov.au/recordkeeping/control/rkms/contents.html.
[42] Available at http://www.archives.sa.gov.au/files/management_standard_metadata.pdf.
[43] Available at http://www.prov.vic.gov.au/vers/standard/.

Testing, cross-validation and revision of the analytical framework (also referred to as the Schema and Analysis and Evaluation Instrument) were conducted by three different analysts who encoded selected archival schemas, some examples of key metadata schemas from related information fields (for example, Dublin Core) and schemas from scientific and artistic domains independently.

Findings about MADRAS tools and instruments

Upon proceeding with analysis of selected schemas, the researchers were somewhat surprised by the spotty nature of schema documentation. Since a schema is analyzed based upon its documentation, it is vital that this information be clear and concise, but often the researchers found it to be insufficient/deficient. Insufficient schema documentation led to the realization that the analysis questions needed to be refined to make sure that they were focused on eliciting responses about what a given schema is intended to do as opposed what that schema "can be made" to do. This in turn led to the realization that very few schemas can be analyzed accurately independent of their implementation.

Although it was agreed that the analysis undertaken within MADRAS should proceed only to the element level, while actually answering the analysis questions, the researchers found that they spent 85% of their time pouring over the definitions of sub-elements. Although this does not necessarily suggest taking analysis down to the sub-element level as a rule, it must be acknowledged that the real meat of a metadata schema does not tend to live at the element level, especially when one is being asked to describe records in the intricate manner proposed by the InterPARES and ISO 23081 instruments.

Because the language used in ISO 23081 and the InterPARES benchmark and baseline requirements differs, it was a challenge to clarify the meaning and intention in each of the documents and then to unify them. This proved to be difficult because the focus of the instruments is quite different. InterPARES focuses on *domain-independent* digital records, while the ISO standard focuses on records in all media made in the *course of business*. In addition, there are times when InterPARES and ISO 23081display such different approaches to a particular recordkeeping problem that the MADRAS analysis questions—in trying to satisfy both "masters"—become confusing. For example: in addressing MADRAS Question 206, "Chronological Date" as opposed to "Creation Date," InterPARES, drawing upon its diplomatics lineage, lists four date types in benchmark requirement A.1.a.iii: "Chronological," "Received," "Archival" and "Transmission." ISO documentation is concerned only with "Creation." Therefore, the picklist for this question, which has to combine the language from both sets of requirements, demonstrates how the combination of two different instruments can cause confusion. In this case, the differences in approach to dates appears to spring from the fact that the InterPARES requirements only admit those dates to which the record keeper can directly attest (i.e., one can identify the date written on a document (Chronological Date) but cannot actually be sure that this was its *creation* date), while ISO appears to believe that the record keeper will be able to identify an authentic creation date.

At other times, ISO 23081 seems overly vague:

> **Example 1**: MADRAS Questions 214 & 215: "Technical characteristics and dependencies of a record" v. "Technical requirements to render or reproduce record"
>
> The ISO documentation makes this distinction, but does not fully explain what makes one different from the other. The researchers assume that "characteristics and dependencies"

is mainly about format, while "requirements to render or reproduce" is more about the entire technical environment needed, but it is unclear.

Example 2: Questions 504 & 507: "Rules that regulate record management" v. "Rules that regulate records management operations"

The ISO Standard is ambiguous. 9.3.1b (which stands behind question 504) states, "capture the business rules or other system controls that regulate record creation and management," while 9.3.1d (which stands behind 506) states, "capture the business rules or other system controls that regulate records management operations." How does the "record creation and management" from question 9.3.1b differ from the "records management operations" of 9.3.1d? The researchers assume that 9.3.1b is about creation, access and use while 9.3.1d is about activities performed only by records managers, such as preservation actions. Furthermore, since these instruments also largely directed how the researchers crafted the system, some of the concepts in ISO 23081 posed particular challenges. The standard describes that the researchers need to capture information "at record capture" and "after record capture." This is not a distinction made in the InterPARES requirements. To incorporate the concept into the analysis tool, the researchers considered metadata about a record's "content, context and structure" to be the metadata created "at record capture." Any other metadata that the researchers describe are, thus, by definition "after record capture." This amounts to isolating the metadata that deal directly with recordkeeping/administration, which appears to be in the spirit of ISO 23081.

Findings about the schemas

As mentioned above, the instantiations provided an interesting commentary on the status of metadata schema publication and documentation practices. It raised issues about persistent identification (for example, stability of URLs for schema documentation), standards for schema documentation and standards for their description addressing lack of and inconsistency in metadata to describe schema documentation.

As noted earlier, analyzing every schema identified as relevant was beyond the scope of this Project. However, the researchers did analyze enough from different sectors and of different types to be able to make the following observations:

- Almost no schema analyzed, with the exception of New South Wales, met all the requirements that were relevant to the schema's stated scope. In general, those schemas that are not designed for recordkeeping prove to be less compliant than the others. It is also often the case that the schemas—no matter the domain—fall short in being able to describe how a record/agent/mandate/business process changes "over time."[44]
- Some schemas were never intended to satisfy the kinds of requirements identified in the analytical framework, but nevertheless address some of them.
- Many record creation or preservation implementations may need to employ more than one schema simultaneously or sequentially to document all relevant aspects of their activities (this is even more likely to be the case where a records continuum approach is being used).

[44] Schemas designed for managing geomatics data may provide an exception to this general observation, since time is a key element of any geographic feature.

- Even if a schema were to meet all the requirements, this is unlikely to be the case in specific implementations/application profiles. The process of completing these selective analyses has demonstrated that many metadata schemas cannot effectively be separated from their implementation. Since it was decided that implementation issues would not be considered during analysis, many schemas appear to fall short in certain areas, and one might even fairly say that some of the analysis questions are poorly answered because of this distinction. For example, the ANZLIC standard (Metadata for Spatial Data Directories in Australia and New Zealand) requires (and the eGMS suggests) an implementation concomitant with the schema that notes which encoding schemas are being used within the implemented XML/HTML tags, not within the metadata elements proper. It must be remembered, however, that not only do existing metadata schemas predominantly not meet the necessary recordkeeping requirements, but actual implementations of specific metadata schemas often only use selective metadata elements and often not in standard ways.

The researchers have also identified that there are two major element/sub-element relationships:

- For a number of schemas (for example, the RKMS/Minnesota group and CDWA), the upper-level elements are only “envelopes” for a series of sub-elements. That is to say, the elements take no data values themselves, but serve as a type of header for the sub-elements, and it is these sub-elements that are actually assigned data values.
- For others (such as eGMS), the elements do take data values, and the sub-elements are actually “refinements” to those values.

Another finding is that some tools, especially those outside the more traditional recordkeeping/archival domain, do not fall neatly into some of MADRAS’ classifications. How can the researchers modify MADRAS to account for this?

> **Example:** CEDARS Preservation Metadata
>
> Element obligation value is not designated as “Mandatory,” “Optional” or “Conditional.” Rather, the coding is based on the level of specificity indicated by the element (i.e., the extent to which it may be usefully applied across a wide range of digital materials). Values used in coding include “less significant,” “very significant” and “significant.”

In the above example, therefore, the element coding is assigned based on the types of objects rather than on the function/purpose of the metadata. So what does this mean? It means that it is difficult to compare element obligation encoding values between schemas since the reason the coding is being applied may differ from schema to schema. In other words, the researchers would be comparing apples to oranges. Moreover, the “significance” value is a subjective coding.

Because the MADRAS questions are so heavily weighted towards business process-specific recordkeeping issues, some non-recordkeeping schemas are not fully appreciated for what they can do. Not surprisingly, and perhaps also not a problem for the purpose of MADRAS, the analytical tool has difficulty evaluating aspects of a metadata schema that address aspects such as depth of description or monetary value that are emphasized by schemas in non-recordkeeping domains (for example, CDWA and ANZLIC). Related to this, as might be expected, granularity of content description required to meet the user needs of those specific communities tends to be higher in non-recordkeeping schemas, while the recordkeeping schemas focus more on context description. Finally, it is noted that some non-archival specialists who register their discipline-specific schemas into MADRAS may have difficulty in doing so, as the MADRAS tool is specifically designed to meet the needs of archivists and is therefore expressed using archival terminology that may be unfamiliar to practitioners in other disciplines.[45]

To discuss these issues a little further, ANZLIC is an example of a complex metadata schema that was not designed specifically for recordkeeping purposes; it is all about describing the *content* of a dataset accurately, and, insofar as this is the case, is more or less doomed to perform poorly in a MADRAS analysis that privileges tracking contextual information over time. Of the thirty-six elements analyzed in MADRAS, twelve of them are used to describe the contact information of the custodian of, or source contact for, the dataset. Ten other elements describe the contents of the dataset (for example, its date, physical location, extent and keywords). Beyond content description, seven elements are used to describe the reliability and quality of the dataset (for example, lineage, positional accuracy and update frequency) and three elements describe access information including both format and rights. The remaining four elements consist of the title, a unique identifier, a date for the metadata record and a spot for "anything else." There is nothing in this schema that allows the user to see how this dataset has been used or housed over time.

On a final note, it seems almost impossible for any single-object schema to measure up to ISO 23081's requirement that a recordkeeping system not only track which mandates/agents/business processes are related to which record, but also track the set of mandates/agents/business processes themselves. In fact, what ISO is describing is the complete recordkeeping system, but most schemas are just meant for the record-centric portion of that system. Ultimately, this would be an implementation issue, because most metadata schemas do not assume that they are the only schema being used in a system. One way to address the issue might be to track the mandates separately, manually inserting the appropriate code or link within the system using the schema at hand.

MADRAS products

MADRAS, as an automated tool that facilitates schema analysis as well as serves as a registry of existing and evolving schemas; the analytical framework as a stand-alone tool that is to be incorporated into ISO 23081 but that can be used independent of both MADRAS and the ISO standard to assess current and draft schemas and application profiles; and the evaluative reports on the schemas analyzed by InterPARES researchers all constitute products of this research.

For each schema or set registered, a set of evaluative reports can be generated that: (a) indicate whether the schema meets all, some or none of the InterPARES benchmark and baseline requirements or ISO 23081 metadata requirements (recognizing that users may be interested in addressing either or both sets of requirements), (b) pinpoint in what ways, if any, the schema falls short and (c) provide guidance as to how the schema could be modified or augmented to meet all the relevant requirements.

One additional product that is still in process is the doctoral dissertation of Lori Lindberg, which is examining the implications of this analysis for ISAD(G), ISAAR, EAD and EAC and making specific recommendations for extensions to those descriptive standards and establishment of a new framework within which future development should take place.

---

[45] This was the case, for example, when a geomatics student worked with researchers at UCLA during the Excel spreadsheet development phase of the tool. The student was a metadata expert in the field of geomatics but required very intensive support from the UCLA researchers to fill in the required fields. In the end, a UCLA researcher had to register the schema and results from that process are pending.

Warrant Database

Scope and rationale

Description Cross-domain researchers made a decision early in the InterPARES 2 Project that developing an entire new metadata schema to address InterPARES requirements was neither practical nor likely to be adopted either within the recordkeeping and archives community or those communities within the various focus areas of the Project. There were several factors behind this decision: the difficulty in developing an all-encompassing schema that would work in so many different settings, issues of how to ensure that the schema would be able to continue to evolve after the end of the Project and difficulties in persuading communities (including archival communities) that had already invested in their own metadata frameworks, to adopt one developed by InterPARES. Instead, it was decided that the researchers would develop a way of assessing those schemas already developed by different communities against InterPARES requirements and provide them with feedback about how they could be extended or modified to address recordkeeping issues. The researchers then discussed how they could develop persuasive arguments that might lead those communities to respond to the Description Cross-domain's recommendations. The researchers decided that they needed to understand better what the communities were already saying about metadata and associated issues such as trust, reliability, authenticity, status as original, accuracy, ownership and custodianship, moral rights and preservation; which individuals were regarded as authoritative on these issues; and to what internal or external mandates they might likely respond. Armed with this knowledge, the researchers felt that they would be in a position to address the relevant communities in terms of their own concerns and mandates, if they existed, rather than appearing to impose InterPARES' upon them.

The literary warrant database was built using the method developed by Wendy Duff as part of the Pittsburgh Electronic Records Project. This involved identifying a warrant for a particular course of action based upon such things as legal or other juridical mandate, professional best practices, professional literature and other social sources.[46] In this case, the researchers were particularly interested in identifying literature and other sources that discussed the need for the creation and continued maintenance of description and other metadata supporting the accuracy, reliability, authenticity and preservation of records and other record-like objects.

Working with input from researchers from other InterPARES groups, the researchers conducted a literature review across each focus area to identify how different communities currently perceive and discuss the need for, and role of, metadata in ensuring the creation and preservation of reliable and authentic materials. The researchers designed and set up the Web-based database to capture standardized literary warrant analyses. The software chosen allowed researchers to input remotely into a single database, but little effort was spent on developing a public interface since initially the tool was developed solely to support the researchers. Guidelines were developed for using database and analyzing warrant, and researchers from the Description and other InterPARES groups were trained in their use so that they could input

[46] See Wendy M. Duff, "The Influence of Warrant on the Acceptance and Credibility of the Functional Requirements for Recordkeeping" (Ph.D. dissertation, University of Pittsburgh,1996); Wendy M. Duff (1997), "Warrant and the Definition of Electronic Records: Questions Arising from the Pittsburgh Project," *Archives and Museum Informatics* 11(3-4): 223–231, Available at http://tort.library.utoronto.ca:8080/bitstream/1778/4315/1/Warrent.pdf; and Wendy M. Duff (1997), "Compiling Warrant in Support of the Functional Requirements for Recordkeeping," *Bulletin of the American Society for Information Science* 23(5): 12–13. Available at http://doi.wiley.com/10.1002/bult.60.

materials they encountered during their research activities. Description Cross-domain researchers then analyzed materials for which records had been in the database, thus populating the database. In 2005, it was decided that the warrant analysis database might be a useful product for the public also, and the data it contained was transferred from UCLA to the University of British Columbia and loaded into a new database with a public interface.[47]

Results of warrant analysis

The database now contains 177 records that include not only bibliographic information but also summaries of the major arguments used in support of metadata concerns within different communities that can be referenced when developing presentations, publications and other InterPARES 2 products aimed at those communities. In this database are identified the warrants that fed into the analysis framework and that, additionally, have influenced each other.

News Archives Survey

Although a series of diverse case studies were conducted by InterPARES focus groups that included the gathering of data about metadata on behalf of the Description Cross-domain (discussed below), the Description Cross-domain was presented in 2005 with a unique opportunity to study contemporary thought and practice in a professional area that has changed both rapidly and radically with the development of online interactive, multimedia technologies—the news industry and its archives. Researchers decided that a survey of perceptions and practices relating to metadata in this industry would provide important insight into how one specific community is addressing metadata and preservation issues more broadly.

Conducting the survey

In recent years there has been a growing awareness that historic news archives in electronic formats are at risk.[48] In the popular media, printed newspapers are frequently described as a threatened species in the digital world, and Wall Street has responded accordingly by undervaluing media properties across the board. Efficiencies gained through automation have wiped out traditional "morgues" with their paper clippings and film negatives, and there are fewer archivists to tend to their born-digital avatars. Even microfilm, that reliable, long-lived preservation medium, is under serious threat from publishers who no longer see the need for it amid a nightly river of page PDFs extracted from sophisticated pagination systems.[49]

In spite of the myriads of information channels available in the Digital Age, newspapers are still cited by historians as the most often used and most important resources in their research.[50] But even as the Library of Congress, with its National Digital Newspaper Project, pursues filming and digitizing 19th century editions, tomorrow night's all-digital output is every bit as threatened as a crumbling volume of newsprint, because the industry and profession are unprepared to handle

[47] Available at http://www.interpares.org/ip2/ip2_warrant_db.cfm.
[48] Victoria McCargar (2005), "Following the Trail of the Disappearing Data," *The Seybold Report* 4(21): 7–14.
[49] Bernard F. Reilly, Jr., "Knowledge Biodiversity: The Perilous Economic of World News Heritage Materials," in *Proceedings of the ACRL Twelfth National Conference, April 7-10, 2005, Minneapolis, MN* (Minneapolis, MN: Association of College and Research Libraries, 2005), 238–243. Available at http://www.ala.org/ala/acrl/acrlevents/reilly05.pdf.
[50] Helen R. Tibbo (2003), "Primarily History in America: How U.S. Historians Search for Primary Materials at the Dawn of the Digital Age," *American Archivist* 66(1): 9–50.

it. Moreover, news is increasingly being created and transmitted to the newspapers from reporters in the field using online transmission of digital text, photographs and video.

Victoria McCargar, an InterPARES researcher and leading authority on electronic news archives, with the assistance of Shannon Supple, at the time a graduate researcher at UCLA, created a survey instrument to benchmark current trends in digital preservation among news archivists.[51] After receiving the appropriate permissions for human-subjects testing through UCLA, the survey was uploaded to a professional interface at the SurveyMonkey Web site in August, 2005. The invitation to participate in the survey was communicated through a popular and very active listserv mounted by the News Division of the Special Libraries Association, which numbers more than 650 news librarians and archivists. The survey was available to participants through the end of October, 2005.

The survey consisted of eighty questions divided into the following categories:

- Institutional environment
- Professionalism
- Budget
- Use of archives
- Policy
- Technology
- Metadata
- Digital preservation
- Copyright

Additional sections allowed for comments and for survey-takers to volunteer contact information if they were willing to participate in follow-up data-gathering. The survey instrument was designed in its initial questions to discover areas in common among organizations, such as which departments have responsibility for archival systems and how archival systems are budgeted. Later questions homed in on issues specific to digital preservation.

Data analysis was begun in February with the goal of making a "first-cut" presentation at the Special Library Association's 2006 annual conference in Baltimore.[52]

Discussion of the survey results

Despite the advances in digital preservation research in the last ten years, there is still a remarkably low level of awareness of the risks to cultural heritage material in the private sector, which falls outside the domains of academic libraries, archives and government. One of the challenges in mounting a preservation survey of news archives was the lack of basic understanding of the issues among potential participants. The goal of the survey instrument was to capture as much data as the researchers could from each participant before she or he reached questions that could not be answered without a fuller understanding of the complexities of these issues. In fact, of the seventy-seven participants who started the survey, only twenty-eight—fewer than half—completed it. Those who did, however, helped paint a picture of a great volume of historic, cultural heritage material at risk.

[51] This "News Archive Survey Instrument" is available on the InterPARES Web site and on the DVD accompanying this book.
[52] See Victoria McCargar (2006), "You Can Kiss Your Assets Goodbye: The State of News Archives at the Dawn of the Digital Era," paper presented at the Special Libraries Association Annual Conference 2006, June 10-15, 2006, Baltimore, MD, USA. PowerPoint slides available at http://www.ibiblio.org/slanews/conferences/sla2006/presentations/assets.pdf. Audio available at http://www.ibiblio.org/slanews/conferences/sla2006/audio/vicky.mp3.

A somewhat more subtle function of the survey was to try to educate survey-takers about digital preservation on a basic level. The question, “How knowledgeable is your staff about digital preservation?” revealed a low level of understanding; 55% of respondents answered “Low,” and almost a fourth stated they had no idea what level of understanding prevailed. Only 15% indicated they had some knowledge, and only two respondents indicated that they had a “high” level of understanding. Questions like this are useful for establishing a benchmark for gauging increasing awareness.

One of the most interesting—and unsettling—questions addressed instances of actual loss: “In any of your previous preservation activities (including upgrading software, moving to a new storage medium, moving to a new software product), did you experience any loss of data or metadata, or otherwise compromise the archives?” Of the twenty-eight responses, only five answered that they had not. Twenty-one of the remaining twenty-three reported some form of loss, ranging from minor (a few corrupt images on CD-ROMs) to the serious (the loss of controlled vocabulary terms for certain objects) to the disastrous (loss of an entire collection of thousands of photographs). The two instances of “don’t know” were telling insofar as they point to an archives environment where data validation is not routine. Indeed, these instances of loss seem to have been uncovered by accident, in the course of a system upgrade or on the fly. If losses are not detected quickly, the chance of retrieving an intact original from backup is lost.[53] Moreover, this lack of routine bit-level validation has implications for data authenticity even in the short term, as will be noted below.

Some of the other results of interest were:

- A low level of commitment by management to archival policy. Only 33% of responding newspapers enjoyed “very committed” oversight. In a future survey it would be worthwhile to explore the extent to which this is a result of revenue interests (mounting the Web sites via archival data feeds) or a commitment to preservation for its own sake.
- The concept of *authenticity* in the digital environment is still rooted in the old model of microfilm as juridical version. To the extent to which news archivists answered that authenticity was a consideration in their archives—about half indicated that it was “important” or “very important—*authenticity* refers to how closely the material in the database reflects what was printed on paper. Bit-level authentication of individual files in the digital preservation sense is an unknown concept. Saying that, larger newspapers do recognize the legal implications of having an “authentic” representation of a printed article or photograph, and some, such as the *Atlanta Journal-Constitution*, have a notary public on the newsroom staff who can validate printed copies from microfilm to fulfil a legal request, either one arising from the newspaper’s own activities or those between third parties.[54]
- A lack of dedicated funding. About 20% of responding news libraries indicated that they had a budget earmarked specifically for preservation, and another 10% had a separate preservation budget. However, it is highly unlikely that this funding factors in digital preservation; it is almost certainly dedicated to *digitization* projects to unlock the commercial value of historic photography, and, ironically, sets up a new preservation problem for the collection of newly scanned JPEGs.

[53] Victoria McCargar (2006), “The Heart of Darkness: A Foray into Aging JPEGs,” *The Seybold Report* 5(22): 9–12.
[54] Personal conversation between Victoria McCargar and Virginia Everett, news director of the *Atlanta Journal-Constitution,* May 9, 2006, in Atlanta, GA, USA.

- A lack of control over the technology environments in which news archivists operate. Only 13% stated that the archivists were responsible for software and 5% for hardware support. In both cases, the responsibility fell to the information technology department and/or the newspaper's vendors. In some instances the photography department was the responsible group.[55] All of these point to a situation where those best equipped to deal with digital preservation—information professionals—are not the major stakeholders in the archives.
- Metadata standards are soft or nonexistent. The reigning schema, IPTC, is widely used (it is the basis for most commercial systems), but of the 58% of respondents who said they use it, up to two-thirds reported that the schema is "somewhat to highly customized" in their archives. The remaining respondents indicated no standard schema or did not know whether one was in place. Schemas associated with digital preservation like PREMIS and MIX (and their envelope METS) are unknown in news libraries.
- There is a proliferation of file formats such as digital video, information graphics, GPS databases and the Web pages in many of the archives as the impact of multimedia publishing matures. However, few controls are in place. More than three-quarters (79%) of responding news libraries reported no policy for handling digital materials over the long term. Of the 21% that have such a policy, only 12% attempted to address problematic, fragile formats, and none of the archivists reported regular reviews to address technological change.
- Similarly, most newspapers do not attempt to capture metadata about these formats, which is considered critically important information in the PREMIS schema. Fewer than 40% of survey respondents indicated that they attempt to catalogue hardware and software metadata in their archives, while only 15% record the operating system and 7% record the necessary peripherals even though all of these elements are specified in PREMIS.[56] These numbers cannot be extrapolated across all news archives since only twenty-six respondents of the original seventy-seven were still participating at this point in the survey and probably represent just the small portion of the community that actually understand digital preservation issues.
- The one area of digital preservation metadata where newspapers are arguably quite thorough is copyright. The U.S. Supreme Court decision in *Tasini*[57] led to the removal of entire sections of many publications, and, in the interim, most papers have better controls in place to identify authorship, ownership and certain aspects of provenance. However, news archivists are much less informed about legal issues relating to preservation of copyrighted material in their digital archives, including reformatting, migrating or normalization. 60% of respondents answered "don't know" when asked about what actions they are legally allowed to take. The remaining respondents who did indicate an awareness of legal issues, were, in many cases, misinformed. Working in units of for-profit institutions, news archivists face proscriptions on preservation activities that are not encountered by nonprofit and public repositories; this is an evolving situation as the

[55] Photographers' archiving practices are highly idiosyncratic; see Jessica Bushey and Marta Braun (2006), "InterPARES 2 Project - General Study 07 Final Report: Survey of Recordkeeping Practices of Photographers using Digital Technology." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs07_final_report.pdf.
[56] McCargar was a member of the PREMIS Working Group in 2004-05 and catalogued a typical newspaper complex/compound object using the draft schema: http://www.oclc.org/research/projects/pmwg/premis-examples.pdf.
[57] *New York Times Co., Inc., et al. v. Tasini et al.* (00-201) 533 U.S. 483 (2001) 206 F.3d 161, affirmed. Available at http://supct.law.cornell.edu/supct/html/00-201.ZS.html.

Library of Congress tackles revisions to the Digital Millennium Copyright Act, the so-called Section 108 Study Group.[58]

Conclusions

In aggregate, the data describes a wealth of historic material in risky, proprietary formats and an important segment of the archivist profession that is ill-equipped to handle them.

Measuring awareness and institutional change over the next few years is important to understanding whether news properties, left to their own devices, will be capable of sustaining this content into the future. News librarians and archivists—practitioners often wear both hats—are well aware that they are responsible for their publications' writing of daily history. The opportunity to comment at the end of the survey questions afforded a few participants a chance to vent their frustration: "The researchers are so busy creating digital archives the researchers are not paying attention to the problems the researchers will leave behind," and "The archival aspect of a newsroom library is often considered an ancillary function of the newsgathering operation, not a key strategic priority for the company." Newspapers, increasingly pressed to boost revenue as advertising shrinks, have hard priorities that may not coincide with preservation; as one survey respondent put it, "In pursuit of the bottom line, management seems to feel that it is more important to spend money than getting the paper out today than it is to archive for the future."

Benchmarking news archives at this juncture will help digital preservationists monitor what might be identified as an impending crisis. But those hoping for solutions to arrive from stronger standards and best practices may be in for a long wait; pursuing a third-party repository model may be a more promising avenue.[59]

Metadata Specification Model

The premise underlying the work of the Description Cross-domain is that detailed trustworthy metadata are key to ensuring the creation of reliable, and preservation of authentic, records and other entities in electronic systems. This argues for an end-to-end metadata management regime that addresses which metadata need to be created and/or carried forward in time, for what purposes, by whom, and how they are to be preserved and validated. Bound up with this, however, are difficult issues associated with how to create rich metadata in a resource-efficient manner as well as how to manage and continue to ensure the trustworthiness of the volume of metadata one ends up accumulating over time (including metadata associated with the preservation, reproduction and dissemination aspects of the archival function). This raises interesting questions such as whether certain metadata can be efficiently segregated and eliminated after validation, certification and summarization by a preserver. Without addressing this question, preservers will ultimately end up managing more metadata than the entities to which they refer.

One goal of the metadata specification model was to identify an overall set of metadata requirements that specify what metadata need to be created, from which sources, how and by

[58] See http://www.loc.gov/section108/. McCargar contributed a public comment on behalf of news archivists (see Victoria McCargar and Peter F. Johnson (2006), "Comments to Section 108 Study Group: News Archives," submitted April 28, 2006. Available at http://www.loc.gov/section108/docs/McCargar.pdf).

[59] McCargar is consulting on a project to develop an audit instrument for a trusted news repository at the Center for Research Libraries; for a brief overview, see Center for Research Libraries, "Auditing and Certification of Digital Archives." Available at http://www.crl.edu/content.asp?l1=13&l2=58&l3=162.

whom, at which points within both the Chain of Preservation (lifecycle) and the Business-driven Recordkeeping (records continuum) models being developed by the IP2 Modeling Cross-domain and retention periods for such metadata. This metadata specification model could then form the basis for developing specifications for automated tools that can be used to assist with the creation, capture, management and preservation of essential metadata for active and preserved records. A second goal was to develop an economical and consistent way of talking about different classes of metadata to facilitate systems design, task allocation and management, as well as automated metadata creation.

Actions taken and products created

Description Cross-domain researchers had to wait until work was sufficiently advanced on the InterPARES 2 activity models to begin work on the development of metadata specification models for the Chain of Preservation and Business-driven Recordkeeping models. Because the former was the more complete toward the end of the Project, the researchers were able to develop a metadata specification model for it.[60] In the metadata specification model for the Chain of Preservation model, the following definition was used for "metadata:" *a machine or human-readable assertion about a resource relating to records and their resources*. Descriptive metadata are defined as those categories of metadata carried forward to be used as evidence for archival description. One hundred thirty-seven (137) different metadata assertions were identified (i.e., different instances of types of metadata), sixteen types of assertions were identified. Two cut across all stages of the lifecycle, one cut across two stages, and the other fifteen were evidenced only in one stage. The resulting model is still a theoretical model that is awaiting validation through instantiation—both through walkthroughs based on the case studies conducted by the InterPARES 2 focus groups[61] of specific implementations and by actual system building. When researchers start to work on the development of the metadata specification model for the Business-driven Recordkeeping model, it is anticipated that the researchers will encounter some of the same issues as were encountered in developing MADRAS in that the records continuum has a very different set of entity foci to the records-centric notion underlying the lifecycle. Other work that is continuing includes the development of attribute pairs for the metadata identified in these models which would designate the values different assertions should take; the development of a typology of classes or categories of metadata and, potentially, the mapping of both metadata specification models onto the OAIS model.

[60] For a general overview of the metadata elements identified in relation to record-making, recordkeeping and preservation activities, see the narrative discussion of the Chain of Preservation Model in the Modeling Cross-domain Task Force Report. For a more detailed description of the Chain of Preservation metadata specification model, see Joseph T. Tennis (2008), "Metadata in the Chain of Preservation Model," *Archivaria* (in press).

[61] For discussions of two walkthroughs done of earlier drafts of the Chain of Preservation Model, see William Underwood, Kevin Glick and Mark Wolfe (2007), "InterPARES 2 Project - General Study 12 Final Report: Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs12_final_report.pdf; and Randy Preston (2004), "InterPARES 2 Project - Modeling Cross-domain: Walkthrough of the Manage Chain of Preservation Model Using Case Study 14 Data," draft report. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs14_COP_model_walkthrough.pdf.

Case and General Studies Data Analysis

Actions taken

One major component of the work of InterPARES 2 was a series of case and general studies undertaken by the arts, science and government focus task forces, examining dynamic, interactive and experiential environments in each focus.[62] These studies examined many facets of these environments and included several questions in their protocols that potentially addressed issues of concern to the Description Cross-domain:

- How are the digital entities identified (e.g., is there a [persistent] unique identifier)?
- From what application do the record system(s) inherit or capture all digital entities and the related metadata (e.g., e-mail, tracking systems, workflow system, office system, databases, etc.)?
- Does the recordkeeping system provide ready access to all relevant digital entities and related metadata?
- Does the recordkeeping system document all actions/transactions that take place in the system re: the digital entities? If so, what are the metadata captured?
- What descriptive or other metadata schemas or standards are currently being used in the creation, maintenance and use of the recordkeeping system or environment being studied?
- What is the source of these metadata (institutional convention, professional body, international standard, individual practice, etc.)?

Description Cross-domain researchers recognized that metadata issues could also surface in a more general manner in the course of the case study.

Case study data analysis

The Description Cross-domain researchers sought to identify, through the answers to the above metadata-specific questions and the data collected overall in the case studies, which, if any, metadata schemas and sets were currently being implemented; whether these schemas and sets were home-grown for this particular creator, required by the software implementation used, native to the creator's sector or discipline and/or a recognized industry or national/international standards; whether or not any metadata used addressed recordkeeping concerns, either fully or in part; and the extent to which real-world implementation of metadata measured up to or surpassed the ideal of the metadata requirements delineated in the Analytical Framework. The following discussion is drawn from the data and reports generated by the case study researchers.[63]

In the arts focus, where it is usually the product of the artistic activity that is the object of concern rather than a record that is the by-product of that activity, only two case studies uncovered use of metadata standards, and none of these were standards developed specifically for recordkeeping, archival or preservation functions. Case study 09(03),[64] the Commercial Film Studio component of the multi-component Digital Moving Images case study, cites use of several common bibliographic description and resource discovery metadata schemas—

[62] For a detailed synopsis of the InterPARES 2 case studies, see the section in the Domain 1 Task Force Report titled "Characterization of the Case Studies." Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_2_domain1_task_force.pdf.
[63] See Appendix 18 for a comprehensive summary of the case study data relating to metadata in each of the focus groups.
[64] See James Turner, et al. (2004), "InterPARES 2 Project - Case Study 09(3) Final Report: Digital Moving Images - Commercial Film Studio." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_final_report.pdf.

Categories for the Description of Works of Art (CDWA), the Dublin Core (DC), the Thesaurus for Graphic Materials I: Subject Terms (TGMI), the Thesaurus for Graphic Materials II: Genre and Physical Characteristics Terms (TGM II) and the Anglo-American Cataloguing Rules, (AACR). Case study 09(04) (WGBH)[65] also cites use of Dublin Core and LCSH, as well as an industry schema, the Public Broadcasting Core (PBCore). Case study 03, the *HorizonZero* case study,[66] uses the CanCore standard, which is derived from the Dublin Core metadata set and is based on and fully compatible with the IEEE Learning Object Metadata (LOM) standard and the IMS Learning Resource Metadata specification. In terms of overall metadata implementation, none of the arts focus case studies indicated conscious attempts to apply metadata, beyond a few efforts to establish file naming conventions, largely for retrieval purposes, some version control and, in some cases, rudimentary tracking of file check-in or -out or file archiving.

In the science focus, records need to be not only reliable and authentic, but *accurate*. Data quality parameters are essential in the sciences. Relevant and sufficient metadata, therefore, need to be created to document data lineage,[67] especially in situations where datasets, models, software applications, datasets or multimedia objects have been acquired from elsewhere, which is often the case particularly in data portals. Scientists will not trust data they access from other sources without metadata that clearly indicate the reliability, authenticity and accuracy of those data.

Several science focus case studies exhibit the use of metadata schemas that, although not originating in the domains of recordkeeping, preservation or archives, nevertheless do address to varying degrees the requirements for the long-term management and preservation of authentic digital entities. Although this is encouraging, it does, in so far as each scientific area continues to define its own metadata standards, nevertheless raise concerns about interoperability and extensibility issues related to collaboration and recordkeeping.

Case study 08, the NASA case study,[68] applies naming conventions and incorporates the metadata elements contained in the Planetary Science Data Dictionary, which defines rules for constructing Data Element and Data Object names within the Planetary Data System (PDS), which are NASA institutional and data type specific Planetary Science Metadata. The case study also includes metadata that are associated with a data product (mainly relating to data processing, although there is a required version element and there are also optional data type and description elements, including mission, instrument and instrument type). When restricted areas are accessed, the system logs the user ID, date, time and operation performed.

---

[65] See Mary Ide (2005), “InterPARES 2 Project - Case Study 09(4) Final Report: Digital Moving Images -WGBH Boston.” Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs09-4_final_report.pdf.

[66] See Brent Lee (2004), “InterPARES 2 Project - Case Study 03 Final Report: *HorizonZero*/Zero Horizon Online Magazine and Media Database.” Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs03_final_report.pdf.

[67] In the field of geomatics, lineage means the history of the dataset, the dataset’s pedigree as it changes form, its lifecycle from collection to acquisition, through all the dataset’s stages of conversion, correction and transformations, its parentage. Specifically lineage contains information that describes the source of the observations, data acquisition and compilation methodologies, conversions, transformations, analyses and derivations to which the data have been subjected, and the assumptions and criteria applied at any stage of their life as well as any biases. In fact, lineage is normally the first part of a quality statement since most other data quality elements are affected by lineage. Data producers have documented procedures and quality requirements they have to meet, and lineage is a kind of audit trail to attest to the fact that the producer has met their standards. For the user, lineage provides a dataset its pedigree, to decide on its fitness for use. The “ultimate purpose of lineage is to preserve for future generations the valuable historical data resource. The key to our understanding of the Earth system may lie in the data collected by past generations” (Derek G. Clarke and David M. Clark, “Chapter 2: Lineage,” in *Elements of Spatial Data Quality*, Stephen C. Guptill and Joel L. Morrison, eds. (Oxford: Elsevier Science, 1995), 13–30). Lineage can also be found in a dataset’s associated publications, reports, and technical notes (see Tracey P. Lauriault et al. (2007), “Today’s Data are Part of Tomorrow’s Research: Archival Issues in the Sciences,” *Archivaria* 64 (Fall): 123–179).

[68] See William Underwood (2005), “InterPARES 2 Project - Case Study 08 Final Report: Mars Global Surveyor Data Records in the Planetary Data System.” Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs08_final_report.pdf.

Case study 14, the Archaeological Records in a Geographic Information System case study,[69] indicated the potential through the software used, ArcCatalogue, to create, manage and edit metadata in XML, based on the Federal Geographic Data Committee (FGDC) Content Standards for Digital Geospatial Metadata or the ISO 19115 Geographic Metadata Standard. Such metadata would indicate from what source (for example, publication, repository, Web site or database) the data were retrieved. However, the process for creating and maintaining the digital entities is ad hoc even though the Geographic Information System (GIS) dynamically links geospatial metadata and descriptive attribute data from a wide variety of sources. File naming conventions are used for digital entities and certain aggregations of files can take on an associative identity of their own. Time tagging of georeferenced information is part of the documentation of the processes of creating digital maps, models and georeferenced visualizations. No formal recordkeeping system external to the application being used is applied, and heavy reliance is placed on the creator in terms of getting access to the files.

Case study 19, Preservation and Authentication of Electronic Engineering and Manufacturing Records case study,[70] conducted by the U.S. National Archives and Records Administration and the San Diego Super Computer Center, examined an engineering experiment to test an XML-based archival format for digital model (CAD) records of machined piece-parts used in high-tolerance manufacturing. The intent of the experiment was to preserve not only the geometric specifications of the model but also its semantically encoded metadata, allowing for their examination by reasoning programs for authentication prior to operationalization in computer-aided manufacturing. The experiment used a domain-specific metadata schema, STEP, the Standard for the Exchange of Product Model Data, a comprehensive ISO standard (ISO 10303) that describes how to represent and exchange CAD digital model information. STEP was extended by adding to it metadata elements in the form of logical expressions that enable reasoning over the topological features of the solid model and its functional context. This logical format, including the data model, was transferred into what was termed a new logical preservation format using the OWL Web Ontology Language, an open-source, public domain XML specification of the World Wide Web Consortium (W3C). OWL is a semantic XML format formally recommended by W3C in 2004 for use as a language to represent machine interpretable content when the content needs to be processed by applications rather than just structured for presentation to humans.[71] The case study report outlines specific metadata schemas delineated in the ANSI Y 14.5 Dimensioning and Tolerancing standard, and the use of a metadata cataloguing system (MCAT). Corporate standards were also used for solid-model and drawing metadata.

Case study 26, the MOST Satellite Mission case study,[72] indicates that MOST researchers chose file formats based upon astronomical best practice and then the metadata created were derived from that file format. Digital entities are identified by unique names and an additional set of unique identifiers. The metadata refer to information such as orbital parameters, telemetry information and target image information. The report notes that some of the metadata fields in

[69] See Richard Pearce-Moses, Erin O'Meara and Randy Preston (2004), "InterPARES 2 Project - Case Study 14 Final Report: Archaeological Records in a Geographical Information System: Research in the American Southwest." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs14_final_report.pdf.
[70] See Kenneth Hawkins (2006), "InterPARES 2 Project - Case Study 19 Final Report: Preservation and Authentication of Electronic Engineering and Manufacturing Records." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs19_final_report.pdf.
[71] See http://www.w3.org/TR/owl-features/.
[72] See Bart Ballaux (2005), "InterPARES 2 Project - Case Study 26 Final Report: MOST Satellite Mission - Preservation of Space Telescope Data." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs26_final_report.pdf.

the FITS files are mandatory, due to the file format being used. In general, no metadata standards are used and the MOST researchers have created their own scheme of important descriptive fields that meet the needs of their particular research project. No formal capture system is in place and access is dependent upon the capabilities of Windows Explorer.

No ability to determine whether a file had been altered, how, when and by whom, such as an audit trail, was identified as being built into any system examined in the science focus case studies with the exception of case study 06, the Cybercartographic Atlas of Antarctica case study,[73] which contains a more detailed account of metadata considerations than any of the other science focus case studies. Case study 06 reports that an Authors' Toolkit will eventually allow changes to associated metadata to be tracked. Case study 06 also conforms to ISO 19115 Geomatics Standards and the case study outlines important metadata elements that should be present and where these should be located (although these metadata are primarily cartographic or relate to the nature or behaviour of multimedia objects contained in the Atlas). Because the Atlas acquires data and their associated metadata from other organizations, it incorporates the metadata that accompanies those datasets. Any digital object being incorporated by the Atlas has been peer reviewed and must be described by the creator using the project's metadata standards. Each object is assessed against the Elements of Spatial Data Quality, including lineage, positional accuracy, attribute thematic accuracy, completeness, logical inconsistency, semantic accuracy and temporal information. The case study reports that authenticity in geography is measured in standard metadata as data lineage. Quality measures are dependent on the type of data and their function (for example, the acceptable margin of error for the precise location and size of a particular ice flow to inform tourist ships is smaller than fish counts to inform fisheries and ecological modeling). Other metadata requirements that are followed include the FGDC and/or British Antarctic Survey Directory Interchange Format (DIF) and OGC interoperability specifications.

The Atlas also tracks provenance and rights metadata associated with multimedia objects that have been incorporated into its content, primarily through a citation, caption or link to a bibliography. Linkages between information objects, their functionality and their associated metadata within content modules are described within an XML document. However, there are no unique or persistent identifiers and there is no formal ID lookup system. All versions of any software code used are tracked using Subversion, a source repository system. Documents that accompany the case study include Elements of Geospatial Data Quality, a Multimedia Metadata Discussion document and a List of Standards Adhered to on the Project.[74]

In the government focus, case study 05, the Archives of Ontario Web Exhibits case study,[75] reports that:

> There is no one recordkeeping system for records generated in the creation of exhibits. Different contributors (most notably the curator, webmaster, scanning technician and manager) each create and maintain their own records of this process. The Web site component files exist on both the development and production servers only. Thus there is no common classification scheme or file naming convention.[76]

[73] See Tracey P. Lauriault and Yvette Hackett (2005), "InterPARES 2 Project - Case Study 06 Final Report: Cybercartographic Atlas of Antarctica." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs06_final_report.pdf.
[74] Additional discussion of metadata aspects of the Cybercartographic Atlas of Antarctica can be found in Zhou, "Profiling and Visualizing Metadata for Multimedia Information in a Geospatial Portal," op. cit.
[75] See Jim Suderman et al. (2004), "InterPARES 2 Project - Case Study 05 Final Report: Archives of Ontario Web Exhibits." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs05_final_report.pdf.
[76] Ibid., 11.

As a result, "[r]ecordkeeping throughout the creation process of a Web exhibit is ad hoc and at the discretion of the participating individuals."[77] Since there is no recordkeeping system for the exhibits themselves, "these Web sites and their contents need to be seen within the context of a corporate Web site which has some aspects of a recordkeeping system."[78] Maintenance of the exhibits primarily involves making revisions to them. However, changes made to the exhibits are not documented. Web logging software documents aspects of all interactions with the institution's Web site, but related metadata are not readily accessible, even if it has been captured.

Case study 18, the Alsace-Moselle Land Registry (AMALFI) case study,[79] examined how Web-based applications enable the creation and management of the ordinances and access to the content of the land registry. No metadata schemas or standards are used, and metadata are not discussed in the final case study report. However, every inscription in the database is numbered with a persistent, unique identifier and dated and there are naming conventions for the other entities. Ordinances are also numbered and dated. Each scanned image of the registers is numbered according to the system already in place for numbering individual pages of the registers. Each inscription in the land registry is also connected to a physical file, the annex, by means of a reference number.

The database aggregates the data according to the main categories: parcels, persons, rights and obligations. The database has been organized following a data model closely mapped on the organization of a single inscription (*feuillet*) within the paper register. That is, the main entity is the inscription, of which there is one for each landowner. Each inscription may hold multiple land parcels and multiple inscriptions within the administrative scope of a land registry office. That is, a single inscription contains information relative to all the properties of a single person within a given administrative territory (usually a commune or part of one).

The system keeps track of changes to the digital records. The relevant fields of the database are updated with the information contained in the ordinance once the latter is signed by a judge. Also, since each land parcel listed in the registry also references an entry in the cadastral survey, any change to the cadastral survey must first be reflected in the land registry. However, scanned images of the paper registry and digitally signed ordinances are never modified.

Case study 20, the Revenue On-line Service (ROS) of Ireland case study,[80] reports on a system whereby taxpayers can pay their taxes online. The system keeps track of changes, which are noted and logged with a time/date stamp and the name of the employee making the change. Metadata issues are only touched upon in the case study report, but appear to be addressed by several "in-house" concepts, including by the capture of what is known as the "security wrapper." The "security wrapper" is "the entire transaction dataset received from the customer by ROS. This includes the transaction element; i.e., tax return and payment instruction, as well as the 'security packaging' element; i.e., digital signature, date/time stamp etc."[81] Another possible use of metadata may be in data transfer, although this is unclear. The report does state that "metadata related to the expired certificates, *in addition to the security wrapper*, is maintained

[77] Ibid., 30.
[78] Ibid., 11.
[79] See Jean-François Blanchette, François Banat-Berger and Geneviève Shepherd (2004), "InterPARES 2 Project - Case Study 18 Final Report: Computerization of Alsace-Moselle's Land Registry." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs18_final_report.pdf.
[80] See John McDonough, Ken Hannigan and Tom Quinlan (2005), "InterPARES 2 Project - Case Study 20 Final Report: Revenue On-Line Service (ROS)." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs20_final_report.pdf.
[81] Ibid., 27.

within ROS."[82] Although an Irish Public Service Metadata standard exists, it is not used with ROS. Nonetheless, "[t]wenty-two schemas for the tax forms available via ROS are publicly available in XML-DTDs for inclusion in ROS-compatible software developed by third parties. Each schema includes a DTD and element definitions and explanations."[83] The standards for these schema are based on institutional practice. "From an operational perspective, form element selection and management are, in large measure, based on data flow and format requirements of the ITP and related back-end systems applications."[84]

Case study 21, the Singapore Supreme Court Electronic Filing System (EFS) case study,[85] reports that the Bankruptcy Section of the Supreme Court has created an internal procedure manual and workflow chart on the process of filing bankruptcy petitions in accordance with juridical requirements. In addition, the Bankruptcy Act (Commencement) Notification of 1995 details the necessary documentary forms of records related to bankruptcy proceedings. There is also a prescribed documentary template allowing law firms to enter information on their cases. A unique, persistent identifier—the file reference number—is assigned to each case. The digital certificates issued and managed by the system have a unique Certificate Control Number. The naming conventions of the records created under EFS are clearly stated under the Bankruptcy Rules and Act as well as in the registry's internal workflow. To organize records, there exists a uniform classification scheme comprising all Supreme Court cases. "The internal business processes and the juridical regulations laid down by the courts govern the organization of the digital entities of the EFS."[86] To make organization easier and more intuitive in the electronic system, "[t]he file classification of bankruptcy records in EFS mirrors its previous paper based filing system, with some modifications. In the traditional paper environment, the record profile of the case file comprised the case number and the name of the debtor. However, the EFS bankruptcy case file comprises not only [these elements], but also the name of petitioner, case status (pending or concluded) and the bankruptcy status (bankruptcy order, adjourned or withdrawn)."[87]

The term "metadata" is rarely used in the final EFS case study report. Instead, it is sometimes referred to as a "prescribed set of information" or, most frequently, a "documentary template," which acts "as the record profile." "The front-end module allows law firms to enter relevant metadata elements using a prescribed documentary template that are in HTML pages and to attach the corresponding supporting records, which are in PDF."[88] Metadata elements that the law firm must enter include the firm's file reference number; party details, including the party type (i.e., whether the firm is representing the creditor or debtor), the name of the parties, addresses of the parties and the name of the solicitor. "The fields of the documentary template are controlled, to ensure consistency and accuracy of information and this explains why there is a drop down menu for some of the data elements."[89] The schemas for the documentary templates are based on the workflow and juridical requirements of the court.

With regards to capturing the digital entities, it is assumed that when the final report speaks of submissions by law firms to the Court, that this process is equated with the EFS capturing the

[82] Ibid., 55. Emphasis added.
[83] Ibid., 68.
[84] Ibid., 69.
[85] See Elaine Goh (2005), "InterPARES 2 Project - Case Study 21 Final Report: The Electronic Filing System (EFS) of the Supreme Court of Singapore." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs21_final_report.pdf.
[86] Ibid., 19.
[87] Ibid., 10.
[88] Ibid., 20.
[89] Ibid., 30.

submissions/records in question. The final case study report notes that "the EFS captures both the metadata of the record and the actual record itself."[90] It is assumed that the "record" referred to here is the PDF document, while the metadata are what are entered via the Web-based application using what is referred to in the final report as a "prescribed documentary template"[91] coded in HTML.

There is a tracking function for actions/transactions made in the system. In addition, the EFS maintains a transaction log, a financial audit log and a violation log. The transaction log maintains all changes to the digital entities in the system such as changes to documentary templates and deletion of records and annotations. The financial audit log maintains changes made to the payment of fees made to the court, while the violation log keeps all changes to the digital entities in the system, such as changes to templates and deletion of documents. The violation log also keeps track of unsuccessful (and potentially malicious) attempts to use functions.

The Supreme Court Registry is responsible for the processing, registration and custody of records. The court's workflow and record keeping systems operate together using Visual Basic, Oracle database and FileNet document management systems. The court uses FileNet, a document management system that indexes and stores the PDF files sent by the law firms. "The court's application system manages all incoming submissions by the law firms as well as outgoing replies by the court."[92] The system includes the "record register," which is essentially an index of documents within the case file. In the traditional paper environment, the register includes the record profile of the various types of documents related to the case, the document number and date the documents were filed. In EFS, the record register exists in the form of a sub-directory. Compared to the paper based system, the EFS record register has an additional record profile: the originator of the document (the person who created the record).

In case study 24, the VanMap case study, which looks at a Web-based map system for the City of Vancouver, the HTML and CFML pages and embedded GIF images are identified by unique URLs.[93] The data fields, layers and groups are also identified by field names, layer names and group names, respectively. Metadata is assigned based on what the VanMap Team thinks would be most useful for users. Metadata generated automatically upon creation of the data have not yet been investigated. "Fortunately, the VanMap Web site includes data sheets listing, at varying levels of detail, the types of data, their origins and the means by which they are included in VanMap."[94] The homepage of the staff edition Web site includes links to the data sheets. The data sheets, which can also be reached from the VanMap toolbar, contain information about layers, layer groups, reports and functionalities. Links to the departments responsible for the data are also provided. There is as yet no classification scheme applied to the City's electronic records. In fact, the classification scheme to be applied to paper records is still under development.

Case study 25, the Legacoop of Bologna Web site, reports that document creation and maintenance procedures regarding the Web site are not documented.[95] The case study reports

[90] Ibid., 23.
[91] Ibid., 20.
[92] Ibid.
[93] See Evelyn McLellan (2005), "InterPARES 2 Project - Case Study 24 Final Report: City of Vancouver Geographic Information System (VanMap)." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs24_final_report.pdf.
[94] Ibid., 5.
[95] See Mariella Guercio (2004), "InterPARES 2 Project - Case Study 25 Final Report: Legacoop of Bologna Web Site." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs25_final_report.pdf.

that "Internal controls, in terms of the content of what is published, [are] not performed except from what [is] prescribed by professional deontology of [those] responsible for any publication. No other defined audits or controls over documentary production are performed with reference to the digital resources."[96] There is a naming convention for the Web site system that appears to generate a unique identifier. Files are assigned names automatically through the editorial system, according to precise rules. "[F]iles are assigned a progressive identifier according to the category of information they belong to, a process that is transparent to users."[97] The identifier is an incremental number that functions as the primary key in the database system. For traditional (paper) records, the originals are filed in folders organized on a very simple classification scheme. There are some metadata recorded in the creator's registry system for traditional records only. The application provides a profile of the registered incoming and outgoing documents. The following parameters are registered: classification code, recipients, object, date and type of document.

General study data analysis

General study 10, Preservation Practices of Scientific Data Portals, involved a survey of the Web sites of thirty-two (32) broadly defined data services, archives, repositories or catalogues in the sciences.[98] The generic term "portal" refers to these services. The primary purpose of the survey was to collect information about the actual practices, standards, and protocols currently used by these portals and their users to ensure access, accuracy, reliability and authenticity of the data. The choice of the portals considered was based on recommendations from InterPARES 2 researchers who were familiar with and used these in their own research work. The portals selected pertained to different communities of practice in sciences such as health, astronomy, biology, engineering, statistics, genetics, geosciences and ecology, to name a few. This research was not intended to be exhaustive but is an overview that discusses the preservation structures in place, or lack thereof, in the examples surveyed.[99]

Most, but not all, of the data portals include metadata, some are very minimalist and include only header files (e.g., IP2SF4, Cambridge Crystallographic Data Centre), others refer to associated peer review articles, some were designed specifically for that particular data set while others adhere to the metadata standards of their discipline (e.g., IP2SF15, Canadian Geospatial Data Infrastructure), Access Portal or institutions (e.g., IP2SF10, World Data Center for Solar Terrestrial Physics).

General study 10 data portal IP2SF14, the Canadian Institute for Health Information (CIHI), includes fifteen databases and registries of health and healthcare information relating to healthcare services across Canada. At the development stage of each database or system, Steering Committee, Advisory Committee or Expert Working Groups are always established and are responsible for instructing and advising on the overall design and data quality. CIHI's Data Quality Strategy provides a common strategy for assessing data quality across all CIHI databases and registries. The Framework is built upon five criteria of quality, each of which has multiple

---

[96] Ibid., 6.
[97] Ibid., 7.
[98] See Tracey P. Lauriault and Barbara L. Craig (2007), "InterPARES 2 Project - General Study 10 Final Report: Preservation Practices of Scientific Data Portals." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs10_final_report.pdf.
[99] For a brief overview of the general study 10 research, see Tracey P. Lauriault and Barbara Craig (2006), "Do Data Access Portals, Repositories, and Catalogues, Preserve or Archive Geospatial and Science Data?" paper presented at the GeoTech 2006 Conference, 18-21 June 2006, Ottawa, ON, Canada.

dimensions: *accuracy* (how well information within a database reflects what was supposed to be collected—this includes documentation of all data processes), *comparability* (the extent to which a database can be properly integrated within the entire health system at CIHI—this includes identifying how conversions might pose problems for the data as well as maintaining accessible documentation on historical changes to the database), *timeliness* (how easily the storage and documentation of data allows one to understand how timely data or reports are), *usability* (how easily data may be understood and accessed) and *relevance* (incorporates all of the other dimensions to some degree but focuses specifically on value and adaptability). Data elements are developed for each individual database and these serve as metadata that describe information at its lowest (i.e., field) and most concrete level. A number of classification systems are used in collecting and analyzing CIHI information, including ICD-10-CA, Enhanced Canadian version of the 10th revision of the International Statistical Classification of Diseases and Related Health Problems; CCI, Canadian Classification of Health Interventions; and ICF, International Classification of Functioning, Disability and Health. No rights management metadata were identified during the case study.

General study IP2SF19, the National Virtual Observatory, includes extensive discussion of metadata about data collections and services that describe data and computational facilities and their locations and then how to use them. Metadata that describe resources include identity metadata (supplies a name and identifier for the resource), curation metadata (describes who supports the resource and availability information such as version and release date as well as the resource's provenance) and content metadata (describes aspects such as data type, sky coverage and spectral coverage, as well as rights management). This last can be applied to resources at various levels of granularity. Metadata relating to such aspects as calibration, consistency, and level of documentation also provide the basis for data quality assessment. Data quality is both subjective and quantitative, and data collections may have no single data quality metric. Although the completeness and consistency of the resource metadata itself may be a reasonable indicator of the associated resource, this is at best a qualitative measure. Each contributing institution uses its own metadata standards or guidelines.

Conclusions drawn from the case and general studies

Prior to examining the case study data and reports, the Description Cross-domain researchers had anticipated that they would see more, and more rigorous, metadata implementation in the governmental and scientific rather than the artistic sector, where legal requirements for recordkeeping or domain-specific data quality standards, measures and assurances respectively frequently provide warrant for the creation of such metadata. Overall, this did prove to be the case, although it would be useful to conduct further case studies in more areas of the arts and sciences to assess the extent to which these case studies are typical of the wider domains covered by those foci.

Although several of the science focus case studies indicate a strong awareness of the need for metadata and the role they can play in ensuring the accuracy and long-term usability of digital materials that is absent from the arts focus case studies,[100] overall, neither focus exhibited any

[100] In particular, the findings of the science focus case studies 06 and 19 show that highly specialized metadata related to a specific domain, discipline or business activity needs to be captured, maintained and preserved to ensure the preservation of authentic, reliable and trustworthy digital records. The records preserver in case study 19 went considerably further, using international metadata standards for resource description and then extending them using semantic metadata expressly designed to enable powerful new means of preserving authentic digital records independent of proprietary software and hardware.

real consciousness of the overall role of *recordkeeping* as opposed to resource identification or discovery or other types of metadata in its activities. The science focus case studies indicate that a rich level of metadata are created or could be created and that there is clearly an overall concern with metadata quality elements as well as for the importance of standardized naming conventions and version control. This is understandable, since metadata are essential for the dissemination of scientific data. In fact, without metadata that support effective data linkage, quality assessment and dataset authentication, scientific data sets have little, if any, long-term value. In addition to the authenticity of datasets, which is linked to a clear lineage recorded in the accumulating metadata surrounding scientific data, the value of data quality and lineage metadata in the sciences is considered axiomatic in that datasets, and the databases with which they interface, have little to no value unless the auxiliary information required to understand and use them correctly—i.e., the metadata—is included in, or inextricably linked to, the datasets.[101] However, scientific metadata standards need to explicitly address archival and preservation requirements as well as data quality and lineage requirements. An additional concern that was raised in the science focus is that despite element-rich, complex metadata schemas being developed in areas such as the geospatial domain, there is, in many cases, little incentive and/or few resources available actually to create metadata.[102] If archival researchers wish to influence these communities and persuade them to add even more elements to their schemas, then the researchers must be able not only to persuade them that it is in their own best interests, but also to help them create such metadata automatically and transparently. Case study 06 appears to be working in this direction, providing creators with an XML-based Authors' Toolkit.

In the arts focus case studies, whatever metadata-related practices there are tend to be idiosyncratic, ad hoc and at the discretion of individuals working with the system. Any metadata standards being implemented have been developed for resource description, discovery and use purposes and not with a view to ensuring the long-term preservation of authentic materials.

In the government focus, there was clearly more concern for evidential requirements and several of the case studies also raised the question of interfacing between digital and paper systems, although the metadata structures in both are generally not as integrated as is the business process. Although many metadata standards do currently exist within different government jurisdictions, the case studies did not reveal that those were being implemented in most of the systems examined.

Overall Results

The work of the InterPARES 2 Description Cross-domain represents the most sophisticated and comprehensive analysis undertaken to date of the requirements and real-life context for metadata that relate to the establishment of reliability and authenticity, as well as the long-term preservation and potential re-usability of digital materials.

There are two particularly noteworthy products or outcomes of this research. The first is the development of actual tools and specifications that can help individuals and institutions from a range of sectors and interests generate and preserve their digital assets in more thoughtful and effective ways. For example, whether those materials be records or other kinds of digital objects,

[101] National Research Council, Commission on Physical Sciences, Mathematics, and Applications, *Preserving Scientific Data on Our Physical Universe: A New Strategy for Archiving the Nation's Scientific Information Resources* (Washington, DC: National Academies Press, 1995), 31.
[102] Collaborative scientific projects are, however, an exception.

MADRAS can be used to identify ways in which they can be created and maintained in ways that will support their intellectual and physical integrity in and over time (although obviously the imperative is stronger for records associated with high degrees of risk or liability than it is for low risk records or non-record materials). Moreover, the development of the metadata specification model, which aligns closely with the OAIS model, will assist systems developers, as well as creators, managers and cataloguers of digital materials, in coping with what to date has been a highly intractable problem—the high costs (in terms of money, time, expertise and storage) of creating and managing optimal amounts of metadata to ensure maximum integrity and usability of the digital materials to which the metadata relate. The model provides a basis for developing automated tools that can systematically create, gather and manage various types of metadata, as well as identifying more closely what needs to be manually created and also what can be summarized and discarded at certain points.

The second noteworthy outcome, and one of the most interesting aspects of this multi-faceted work, is documenting the many levels upon which metadata work and need to work. The development of MADRAS established an ideal against which existing or draft metadata schemas and sets can be assessed.[103] The assessment conducted by InterPARES 2 researchers of selected schemas indicated that even recordkeeping or archival schemas fall short of that ideal, and non-recordkeeping schemas, as might be expected, fall much further short. However, that analysis also pointed up how the schemas are themselves, within the communities that generated them, ideals and that application profiles vary considerably from implementation to implementation, often stripping down a schema to what are considered to be “essential” elements or the elements that a given system is able to support or the creating institution or individual is able to afford or has sufficient expertise to create. Finally, coming a long way behind all of these considerations, are the actual implementations examined in the focus group case and general studies and the news archive survey, where there was scant evidence (with only a few notable exceptions in the science and government areas), especially in the arts focus, of any attempt to implement recordkeeping metadata at all. Although the trend in information management is toward the creation of leaner metadata, the researchers believe that it is important to contemplate how to change the dynamics of metadata depreciation and minimalization so that they work more in favour of the complexities of recordkeeping and preservation—educating communities and individuals more thoroughly about the role rich and rigorous metadata play in addressing needs that they may not even recognize until it is too late to do anything about it; and developing more specifications that could be built into off-the-shelf as well as customized software.

One major question surfaced by the Description Cross-domain’s work arises not only with the differing scopes and viewpoints of the metadata schemas that have been registered and analyzed by the metadata schema registry, but also in the development of the analytical approach embedded in MADRAS and in the metadata specification models—Should these tools support a single or multiple worldviews on the activities, roles, responsibilities and points of engagement with the record? One of the great contributions and benefits of the InterPARES research over the past several years has been that it has brought together archival researchers not only from academe and practice, but also from very different archival traditions. This, however, has also led to moments of confusion and even contention as the divergent underlying perspectives and practices emerge and must be disambiguated and addressed if they are to be operationalized as

[103] Although it should be noted that as it is, it is difficult to perform a sophisticated interpretation of the analysis when the researchers are holding up all of these very different schemas emanating from very different domains, to a single standard set of questions born of a compromise made from two very different warrants.

tools. The Description Cross-domain researchers found themselves faced with two alternatives—one being the development of research products that tolerate and support more than one approach, the other being to attempt to reconcile approaches that appear at first, and maybe even at second glance, to be irreconcilable.

The Description Cross-domain attempted to straddle both of these alternatives. However, having made a conscious decision to assess the metadata implications of both of the dominant existing models, the relative extensiveness of the Business-driven Recordkeeping Model, with the dimensionality afforded by its four axes of identity, evidentiality, transactionality and recordkeeping entity,[104] necessitated that the Description Cross-domain take a more complex view of metadata and archival description than might have been needed if it had looked only at supporting a Lifecycle Model.

The activity models developed in InterPARES 1 were based on a lifecycle view and presumed a custodial approach to the preservation of archival records. The benchmark and baseline requirements identified responsibilities and capabilities for both the *creator* and the *preserver* but were still predicated upon the physical transfer of records into an archival repository. However, the Description Cross-domain has also had to address the fact that while these two theoretical models currently exist (and it is, of course, quite possible, that further models might emerge in the future), many different kinds of implementations also exist. Some of these implementations adhere to the traditional lifecycle view, but increasingly continuum thinking is influencing practices not only in Australia, but also in Northern Europe and the United States. What is more, archivists and other records keepers who are grappling with the challenges of electronic records, are developing their own hybrids of both approaches. In this context, it should be noted that although, historically, they have been linked closely together, conceptually it is not required that custodialism and non-custodialism be tied to adherence to the lifecycle and continuum worldviews, respectively. It is also important to bear in mind that the world outside of archival science does not use these models, at least not conceived of in these terms, but communities other than archival communities are also targeted user groups for the metadata schema registry and analytical framework and their needs must also be addressed.[105]

These results indicate several deficiencies and challenges in the current state of metadata for these purposes, deficiencies and challenges that need to be tackled by a variety of parties. For the archival and recordkeeping professions, there is a need to acknowledge and address the fact that there are both two worldviews (lifecycle and continuum) that necessitate different metadata specifications, and that the field has also developed two rich sets of metadata requirements that work from very different perspectives and have different degrees of focus and scope (InterPARES and ISO 23089). The work of the Description Cross-domain has also pointed out areas where each set is hard to operationalize in schema and system design. For developers of metadata schemas, MADRAS registration has found all to vary widely in terms of their documentation or meta-information, and some effort to standardize such materials would assist in long-term registration and management of metadata schemas. Moreover, schema developers need to recognize that proprietariness of schemas and their documentation works against schema

[104] Frank Upward (1996), "Structuring the Records Continuum, Part One: Post-custodial Principles and Properties," *Archives and Manuscripts* 24(2): 268–285. Online reprint version available at http://www.sims.monash.edu.au/research/rcrg/publications/recordscontinuum/fupp1.html; Frank Upward (1997), "Structuring the Records Continuum, Part Two: Structuration Theory and Recordkeeping," *Archives and Manuscripts* 25(1): 10–35. Online reprint version available at http://www.sims.monash.edu.au/research/rcrg/publications/recordscontinuum/fupp2.html.

[105] The Open Archival Information System (OAIS) Reference Model is a good example of a high-level model that, at first glance, seems to be a re-expression of a lifecycle model, but upon further scrutiny could equally well support a continuum approach.

registration and long-term management and preservation. Also, MADRAS analysis indicates that almost all schemas, including descriptive schemas developed by archivists in different national and international contexts, also fall short in addressing the needs of electronic records. For systems designers and the builders of automated tools, the metadata modeling demonstrates how more comprehensive metadata could be efficiently created and effectively managed across the life of the system and the records it contains. Finally, for funders of activities that create electronic systems containing records or with the potential to create records, this research raises the question as to whether a metadata creation, management and preservation regime should be required for those systems.

Areas for Future Research and Development

Several areas for future research and development emerged from the work of the Description Cross-domain. Two potential research questions are discussed below:

> *Can metadata associated with the creation and active use of records ever contribute to archival description, particularly in the capture and elucidation of certain kinds of context and fundamental identification and arrangement information relating to the records?*

One aspect of an integrated metadata creation and management regime that makes some in the archival community nervous is the notion, also raised by projects such as the Archivists' Workbench,[106] that certain types of metadata, created while the records to which they relate are active, could be captured or analyzed automatically and used to partially automate or even to replace archival description. As identified by InterPARES 1, records have many types of interacting contexts that need to be documented. Often with electronic records, because of their virtual nature and also their complexity, it can be more difficult to identify these contexts than it might be with traditional records. However, often it is the case that the system within which the record has been created or maintained has in place metadata mechanisms, or could be designed to have them, that document some of the context in which archivists are interested (albeit that these are generally created contemporaneous with the record and lack the hindsight and birds-eye view of the archivist).

Indeed, what is distinctive about recordkeeping metadata is the range of ways in which they can automatically capture salient contexts of records as they move through time, space, systems and types of use and user. For example, metadata can provide detailed descriptions of business processes and logs or audit trails of any changes made to records and associated dates. They can also describe the functionality of the original technical environment and enable users to distinguish the authoritative record from drafts and derivative versions. Metadata can also link separately stored data or record content to the appropriate documentary form to facilitate creating an imitative authentic copy of the original (an approach akin to that being used with the Persistent Archives Technology).

---

[106] This project involves the development of a prototype, infrastructure-independent management tool for software-dependent records in the form of a software application called the "Archivist's Workbench." For more information, see San Diego Supercomputer Center (1999), "Methodologies for the Long-Term Preservation of and Access to Software-Dependent Electronic Records," NHPRC Proposal, June 1, 1999. Available at http://www.sdsc.edu/NARA/Publications/nhprc_latest.pdf. A summary version of the proposal is available at http://www.sdsc.edu/NARA/Publications/nhprc_summary.pdf.

In the future, time and cost concerns as well as new technological capabilities are likely to necessitate that even archival description may be created, at least partially, by automated means, likely including harvesting and re-purposing metadata created by others prior to the records coming into archival custody. For this to be acceptable as an assistance or augmentation to archival description, however, a) the metadata harvested should supplement manual description or should capture some aspect that it is difficult or impossible to do manually; and b) archivists should assess what they do manually in traditional description and identify at the point of recordkeeping systems design what could be captured automatically out of the system. Neither of these activities, however, necessarily usurps the archivist's prerogative to supplement and synthesize the metadata gathered automatically in the process of creating a descriptive instrument. Moreover, because the metadata thus gathered are likely to be in digital form, the archivist would have the option of retaining it both in its original form, as evidence of the records and recordkeeping to which it relates, and to transform it into a form that is more useful for secondary use.

Can metadata-based automated tools support any new kinds of capabilities for the description and use of preserved digital materials?

Recordkeeping metadata are created in a variety of ways and by a variety of agents—they may be created manually (as is the case with most archival description) or automatically (as, for example, would be the case with an inverted index of terms culled from a text document). They may also be automatically inferred, derived or harvested from the records and recordkeeping systems themselves, an approach that looks increasingly attractive as systems developers and information professionals of all types become more aware of the burgeoning overhead of metadata creation and management necessary to support the online provision of trustworthy information. They may even be exploited and re-used for purposes for which they were never intended, such as for corporate knowledge mining, developing new institutional market segments or developing learning objects. In the archival community, research and development activities such as the Archivists' Workbench and PERM Projects of the San Diego Supercomputer Center have begun to explore the development of automated tools for metadata creation and management, as well as for the manipulation of records by end users, and the Clever Recordkeeping Metadata Project identified and prototype innovative ways of multi-purposing harvested recordkeeping metadata.

Approaches such as these potentially not only offer archivists a faster and less labour-intensive way to gain a measure of intellectual control over large volumes of electronic records, but also offer secondary users a much richer set of tools through which to access, manipulate and interpret archival records. They can also potentially support validation mechanisms for recordkeeping metadata and monitor the continued integrity of critical linkages that exist between records and their metadata. Perhaps the most important potential use of automated metadata tools, however, might be to support a metadata management regime, something which, if not automated, would be practically impossible for archivists to implement.

In terms of development work, the researchers hope to revise MADRAS so that it is more usable and useful by communities and researchers who are addressing metadata concerns. This would involve extending MADRAS' content and re-thinking its presentation and outputs. The researchers recognize that in the current incarnation of the reports generated, some of the information entered while registering the elements for each schema (including encoding schemas and repeatability) is not used in the evaluation. An improved report might weight schemas based on such information. For example, if a requirement is satisfied by a required element or sub-element of a given schema, that schema would be designated stronger in that area than a schema that left such requirements to their non-mandatory elements. In addition, the researchers might make use of the presence or lack of an encoding schema specified for an element or the sub-elements of a schema. A schema whose element or sub-element has an encoding schema would be considered more robust than one that does not. One could also see that refining the report to provide the user with an analysis based specifically on how the schema performed within the various recordkeeping entities would be useful. In this way, the user could learn not only the strengths and weaknesses of the schema, but also more clearly where those strengths and weaknesses lie.

Integrating element-description-level information into the analysis and then testing the implications of an element's repeatability or its optional/mandatory status would greatly enhance the analysis of a schema's recordkeeping capability. It would be helpful to increase the amount of analytical information about the encoding schemas required by each schema. The assumption is that there will be times when the analysis can demonstrate that a schema element with specified encoding schema is stronger than one with no encoding schema. This may not be the case for all elements, however. For example, "title" would rarely be made stronger by the use of an encoding schema. Moreover, when registering elements, the Description Cross-domain researchers found that it is rarely the case that a metadata schema *requires* a given encoding schema for a particular data value. This information should be taken into account. Nevertheless, in cases where two recordkeeping schema each have elements covering the five recordkeeping entities (record, agent, mandates, business process and RK process), could the researchers compare these schemas by looking at any encoding schemas which are or are not required for each? Furthermore, does this vary from domain to domain? Would a schema used in the arts domain have different encoding schema requirements for the recordkeeping entities? Encoding schemas facilitate information retrieval, however, and at present MADRAS focuses on issues of metadata creation/preservation. To increase the emphasis on issues such as encodings would suggest an alteration in the focus of the tool. Another approach to increasing the information gleaned from the registration of the metadata elements might be to type the elements into certain categories (content, context and structure, for example) to get a feel for the overall goal of the schema. Then the analysis could take this information into consideration and not judge a description-heavy schema in the same way it does a context-heavy one.

A future iteration of MADRAS should examine whether the ranking of questions should be re-thought and apply that information in the generation of reports. This would require evaluating each question and giving it a weight as well as deciding what element information is absolutely necessary. For example, does a subject classification have more or less weight/importance than, say, the identification of an agent? Another issue for further examination is whether the division of questions by recordkeeping entity actually works well for MADRAS. Automating the analysis tool forced the Task Force researchers to, in effect, make the relationship between the two instruments (and the questions themselves) very rigorous and, as a result, many issues had to be framed as absolutes. In future implementations of MADRAS, the researchers would like to see the reporting become much more sophisticated such that these seemingly cut-and-dried questions could regain much more of their original nuance.

PART SEVEN

STRUCTURING THE RELATIONSHIP BETWEEN RECORDS CREATORS AND PRESERVERS

Policy Cross-domain Task Force Report

by

Luciana Duranti, The University of British Columbia
Jim Suderman, City of Toronto Archives
Malcolm Todd, National Archives of the United Kingdom

Introduction[1]

Policy considerations are all-pervasive in the aims of a research project such as InterPARES because record creation, maintenance and preservation are integral components of many human activities and need the same explicit directions as those for the activities themselves. Policy statements should be most explicit in the juridical dimensions of record authenticity that the archival dimensions must satisfy and in the moral and ethical requirements of records preservation arising from the function of carrying forward the traces of societal memory. Implementation of the findings and recommendations of the two phases of the InterPARES Project are dependent on engagement with the agendas of policymakers to advocate the benefits of implementing the Project's recommendations. This report reflects primarily the research undertaken by the members of the Policy Cross-domain Task Force, especially in relation to legislation relating to the authenticity and maintenance of records and the rights and obligations on creators and users of records.

In practice, all InterPARES areas of enquiry are touched by and have implications for policy at an international, national, sectoral and organizational level. This is the reason why the Project has established in its second phase, hereinafter called InterPARES 2, a dedicated Policy Cross-domain that inherited the function of the Strategy Task Force of its first phase, hereinafter called InterPARES 1, which defined policy as:

> a formal statement of direction or guidance as to how an organization will carry out its mandate, functions or activities, motivated by determined interests or programs.[2]

As such, policy may be expressed through laws, regulations, standards (professional, industry and technical), ethical codes, codes of conduct or practice and guidelines. The implementation of laws and regulations related to the creation and maintenance of records—in other words, public policy—is expressed, or should be, in an organization's records policy. The implementation of a records policy by an organization or an individual, in turn, results in, or should result in, the creation, maintenance and preservation of records, and their associated metadata, that can be used for further action and reference and as evidence of the activities from which they result.[3] InterPARES 2 researchers involved in the Description Cross-domain team explore elsewhere the relationship between records, records' metadata and preservation.

The Policy Cross-domain has examined the policies and strategies that affect the preservation of authentic digital records produced in the course of artistic, scientific and governmental activities. As organizational activities adopt increasingly rich *yet dynamic and thus somewhat unstable* technologies, the preservation challenge grows. The use of dynamic, interactive and experiential systems to carry out organizational activities reflects the common practice of adopting technologies without considering, let alone resolving, the preservation challenges that they present for the records generated and kept in them. Indeed, while this is clearest in e-government policy initiatives, it is also a fact of life in the sciences and the arts.

---

[1] The authors acknowledge the general contribution of all members of the Policy Cross-domain in the preparation of this report. In particular, we thank Mahnaz Ghaznavi, Ken Hawkins and Tracey Lauriault for their contributions to the text and editorial guidance. Any errors of representation or omission are the responsibility of the authors.

[2] Luciana Duranti et al., "Part Four – An Intellectual Framework for Policies, Strategies, and Standards: Strategy Task Force Report," note 1, in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 118. Online reprint available at http://www.interpares.org/book/interpares_book_g_part4.pdf.

[3] See, for example, a recent court case in the United States, *Williams v. Sprint/United Mgmt. Co., 230 F.R.D. 640 (D. Kan. 2005)*, which features the admissibility of metadata as an integral part of legal discovery. Available at http://www.ksd.uscourts.gov/opinions/032200JWLDJW-3333.pdf.

New models for collaboration and production, the outsourcing of activities and functions and the privatization of many parts of the public domain introduce new challenges for records retention. Legislation, case law and multi-national agreements form an intricate and often inconsistent and internally conflicting regulating infrastructure that, rather than facilitating the proper creation and use of digital objects, makes these activities increasingly complex.[4] Taken together, recent changes in technology, public policy and business models have put at risk the ability of organizations to undertake some of the activities necessary for the preservation of records. As a result, part of the Policy Cross-domain research has been to identify and counter specific *barriers* to preservation.

The final product of the Policy Cross-domain consists of an intellectual framework for policy development comprising two sets of principles that distil all the other research activities of the research team. The principles were conceived as instruments to fulfil the research goal of the Cross-domain—to develop model policies and strategies for the long-term preservation of authentic digital records—in the most concise and effective form. By following the appropriate set of principles, records creators or preservers will be able to develop organizational directives, in the form of guidelines, instructions or policy proper, capable of ensuring the continuing preservation of authentic digital records according to methods that reflect and allow for the correct implementation of the findings of the whole InterPARES research. Ideally, these principles should also be enshrined in supranational, national, sectoral or organizational policies, strategies and standards.

Other research units of InterPARES 2 have developed guidelines aimed at achieving the same sort of outcome in a much smaller organization, even down to individual practitioner level.[5]

Research team

Team objective

The InterPARES Project set out the following responsibilities for the Policy research team:

> The Policy Research Team will analyze the existing policies and strategies in each domain and focus of inquiry in light of the work being done by the working groups and then distil from the findings and products of the working groups policies, strategies and guidelines for the reliable and accurate creation and maintenance of the records under examination, and their authentic preservation within the context of each activity and culture generating them.[6]

Throughout the duration of the research, the three focus task forces (i.e., arts, science and government) carried out records creator-based case studies. The data from these case studies and the diplomatic analysis and modeling activities carried out on them are the core research data of InterPARES 2. However, to reach its own specific objectives, the Policy Cross-domain conducted research on international and national legislation, regulations, directives, etc., to

[4] Susan Gutman, Luke Meagher and Adele Torrance (2006), "InterPARES 2 Project - Copyright Policy Annotated Bibliography, Draft version 4." Available at http://www.interpares.org/display_file.cfm?doc=ip2(biblio)_copyright-annotated.pdf.

[5] For the preserver's procedures, see the Domain 3 Task Force Report and the *Preserver Guidelines* in Appendix 21. The Guidelines also are available in booklet form at http://www.interpares.org/ip2/display_file.cfm?doc=ip2(pub)preserver_guidelines_booklet.pdf. For the creator's procedures, see the Domain 1 Task Force Report and the *Creator Guidelines* in Appendix 20. The Guidelines also are available in booklet form at http://www.interpares.org/display_file.cfm?doc=ip2(pub)creator_guidelines_booklet.pdf.

[6] See Luciana Duranti (2001), "International Research on Permanent Authentic Records in Electronic Systems (InterPARES): Experiential, Interactive and Dynamic Records," SSHRC MCRI InterPARES 2 Project Proposal, 412-2001, 1.1-7. Available at http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf.

determine the guidance presently provided to the development of policies and strategies and the issues they raise in relation to long-term preservation of authentic records.

Team composition

The Policy Cross-domain comprised researchers from a mixture of academic, archival and cultural heritage institutions, assisted by graduate research assistants from the universities of British Columbia and California, Los Angeles. During the first half of the Project, the Cross-domain was chaired by Sharon Farb of the University of California, Los Angeles, and Livia Iacovino of Monash University, Australia. In mid-2004, the chairmanship passed to two of the present authors, both of whom were previously team members. Tasks such as data collection and initial analyses were typically carried out by research assistants under the leadership of the researchers, who undertook more involved and complex analyses, wrote reports and liaised with other research units. Some doctoral students also participated in the latter tasks. The Cross-domain's international membership helped overcome language barriers where precision in recognizing the importance of juridical and other policy instruments is of the essence.

The following is a complete list of researchers and research assistants who participated in the Policy Cross-domain Task Force throughout the Project.

Co-chairs:

Sharon Farb	2001-2004
Livia Iacovino	2001-2004
Malcolm Todd	2004-2006
Jim Suderman	2004-2006

Researchers:

Howard Besser	New York University, USA
Hannelore Dekeyser	Katholieke Universiteit Leuven, Belgium
Luciana Duranti	The University of British Columbia, Canada
Philip Eppard	University of Albany, State University of New York, USA
Sharon Farb	University of California, Los Angeles, USA
Mahnaz Ghaznavi	J. Paul Getty Trust, USA
Kevin Glick	Yale University, USA
Elaine Goh	National Archives of Singapore
Maria Guercio	University of Urbino, Italy
Chenhui Hao	State Archives Administration of China
Livia Iacovino	Monash Univerisity, Australia
Terry Maxwell	University of Albany, State University of New York, USA
Evelyn McLellan	Insurance Corporation of British Columbia, Canada
Du Mei	State Archives Administration of China
Shelby Sanett	U.S. National Archives and Records Administration
Jim Suderman	Archives of Ontario, City of Toronto Archives, Canada
Kate Theimer	U.S. National Archives and Records Administration, USA
Malcolm Todd	National Archives of the United Kingdom

Research Assistants:

Barbara Bean	University at Albany, State University of New York, USA
Jessica Bushey	The University of British Columbia, Canada

Natalie Catto	The University of British Columbia, Canada
Erin Coulter	The University of British Columbia, Canada
Seth Dalby	The University of British Columbia, Canada
Jennifer Douglas	The University of British Columbia, Canada
Adam Farrell	The University of British Columbia, Canada
Fiorella Foscarini	The University of British Columbia, Canada
Susan Gutmann	The University of British Columbia, Canada
Peggy Heger	The University of British Columbia, Canada
Sarah Henshaw	The University of British Columbia, Canada
Sharif Khandaker	The University of British Columbia, Canada
Greg Kozak	The University of British Columbia, Canada
Tracey Krause	The University of British Columbia, Canada
Yvonne Loiselle	The University of British Columbia, Canada
Luke Meagher	The University of British Columbia, Canada
Catherine Miller	The University of British Columbia, Canada
Rachel Mills	The University of British Columbia, Canada
Jane Morrison	The University of British Columbia, Canada
Emily O'Neill	The University of British Columbia, Canada
Cara Payne	The University of British Columbia, Canada
Hema Ramasamy	The University of British Columbia, Canada
Geneviève Shepherd	The University of British Columbia, Canada
Melissa Taitano	University of California, Los Angeles, USA
Adele Torrance	The University of British Columbia, Canada
Catherine Yasui	The University of British Columbia, Canada
Sherry Xie	The University of British Columbia, Canada

Research Methodology

The following excerpt from the Policy Cross-domain's research statement describes the team's methodology:

> The Policy Research Team will research and analyze the existing policies, strategies, guidelines and standards in each of the focus areas in relation to each of the domains, examine how they may apply to the digital environments under investigation, compare them to recognize commonalities and differences and identify gaps, especially in relation to the new issues arising from the accessibility, use, manipulability and fragility of the types of records being studied. It will then examine the results of the case studies and of the work carried out in the three domains. On the basis of this analysis, it will articulate principles that should guide the development of policies, strategies and standards for the creation, maintenance, appraisal and preservation of the records in question and give them to the national and multinational teams for contextualization. Upon receiving the requested feedback, the Team will produce guidelines for those responsible for developing policies, strategies and standards at the international, national and organizational level.[7]

[7] InterPARES 2 Policy Cross-domain Methodologies. Available at http://www.interpares.org/ip2/ip2_policy.cfm.

Owing to the time required to carry out the separate rounds of case studies, typically one year to eighteen months, the early phases of the team's research had little case study data with which to work. As a result, the first tasks undertaken were concerned with gathering other sources of existing policy and with their analysis.

Policy of the highest level, such as national and supranational laws and directives, was given priority. Liaison with other research teams, particularly Focus 3 (government), which had several researchers in common with the Policy Team, was particularly helpful in this respect. Key themes were discussed, such as authentication methods, the issues—new for InterPARES in this second phase of the Project—of accuracy and reliability, and emerging technologies such as Digital Asset Management. Policy data were compiled in comparative tables. Individual researchers worked on issue papers or scholarly papers related to their own jurisdiction and interests, while an InterPARES 2 moral rights panel discussed key challenges at the 2004 Conference of the Association of Canadian Archivists in Montreal.

After the mid-term InterPARES plenary workshop of September 2004, with the emergence of the first case study data, five discrete studies, briefly outlined below, were undertaken by the Policy Team.

1. An annotated bibliography on policy related to intellectual property covering a selection of national and supranational jurisdictions[8] and a study of the case study data relevant to the following research questions: "To what extent does society restrict use and impede preservation to protect the interest of copyright holders? To what extent are limitations to copyright being eroded by amendments to existing laws that focus on digital content?"

   The annotated bibliography covers current changes to national legislation in a number of countries, changes that have been introduced as a result of efforts to implement provisions of transnational agreements to which the respective countries are signatories, most notably the World Intellectual Property Organization Copyright Treaty (WIPO WCT). The provisions of this treaty include copyright protection for software as well as digital works and introduce criminal penalties for infringement, which ranges from unauthorized copying of material placed on a Web site to the removal or alteration of rights management controls from digital works. The newly introduced restrictions on re-use are not balanced by adequate exemptions or protections that enable records preservation activities. This precarious balance is further complicated by the trend that sees terms of copyright coverage being extended in most countries by the addition of years, or scope of coverage, or both.

2. A study of policy on privacy and freedom of information policies,[9] which examined the challenge brought to record integrity and authenticity by privacy protection.

   The study was supported by two principal scholarly papers: a comparative regulatory study of Canada, the United States, Australia and the European Union and a second, more

[8] Gutman et al., "Copyright Policy Annotated Bibliography," op. cit.
[9] Malcolm Todd (2005), "InterPARES 2 Project - Policy Cross-domain: Information Policy - Privacy Report." Available at http://www.interpares.org/display_file.cfm?doc=ip2(policy)privacy_report.pdf. The main contributing papers are Livia Iacovino and Malcolm Todd (2007), "The Long-Term Preservation of Identifiable Personal Data: A Comparative Archival Perspective on Privacy Regulatory Models in the European Union, Australia, Canada and the United States," *Archival Science* 7(1): 107–127; and Malcolm Todd (2006), "Power, Identity, Integrity, Authenticity, and the Archives: A Comparative Study of the Application of Archival Methodologies to Contemporary Privacy," *Archivaria* 61 (Spring): 181–214.

theoretical discussion of the issue by triangulating multiple archival viewpoints. Both papers and the summary study propose detailed policy recommendations to promote the preservation of authentic digital records in a way compatible with privacy principles.

3. A study of general records-related legislation, including that enabling archival institutions, evidence acts, etc., from thirteen (13) selected jurisdictions, aimed at identifying commonalities affecting records preservation and potential barriers to preservation.[10]

   The study reviewed national and sub-national legislation as well as the regulatory environment of the European Union. The study examined how records were defined, assessed how comprehensively the records lifecycle was reflected in the rules and looked for consistency (or its lack) in multi-jurisdiction environments.

4. A study of record authenticity.[11]

   The study analyzed the juridical concepts embedded in evidence legislation in the North American, European and Chinese jurisdictions, compared them with the benchmark requirements issued by InterPARES 1[12] and evaluated the digital authentication requirements within the same systems.

5. A study of the potential contribution of open source software and open standards to the long-term preservation of digital records.[13]

   The study examined whether the acquisition policies and transfer procedures of a broad variety of archival institutions showed a coherent body of knowledge on the issues of file format selection generally and the use of open source and open standards specifically. Some highly developed open source policy material was observed in the science Focus data collected in association with case study 06 and general study 10. This is an example of highly specialized usage and high capital cost of unrepeatable data creation forcing the consideration of creation standards from the systems design stage and even in sectoral and funding policies.

The studies were presented at the InterPARES plenary workshop in Chicago one year later. The last study was spun off for completion into the Appraisal and Preservation Domain, Domain 3, as this was deemed a more appropriate research unit for the study. A second Policy panel presented the findings pertaining to the legislation studies at the Association of Canadian Archivists annual conference in June 2006.

[10] Jim Suderman, Fiorella Foscarini and Erin Coulter (2005), "InterPARES 2 Project - Archives Legislation Study Report."Available at http://www.interpares.org/display_file.cfm?doc= ip2(policy)_archives_legislation_report.pdf. Jurisdictions studied are Australia, Canada (including the provincial jurisdictions of Nova Scotia, Quebec, Manitoba, and British Columbia), China, the European Union, France, Hong Kong, Italy, Singapore and the United States. The underlying studies are available on the InterPARES 2 Web site at http://www.interpares.org/ip2/ip2_documents.cfm?cat=policy.

[11] Luciana Duranti (2005), "InterPARES 2 Project - Policy Cross-domain: Authenticity and Authentication in the Law." Available at http://www.interpares.org/display_file.cfm?doc=ip2(policy)authenticity-authentication_law.pdf. The underlying studies are available on the InterPARES 2 Web site at http://www.interpares.org/ip2/ip2_documents.cfm?cat=policy.

[12] Heather MacNeil et al., "Part One – Establishing and Maintaining Trust in Electronic Records: Authenticity Task Force Report," in Duranti, *Long-term Preservation*, op. cit., 19–65. Online reprint available at http://www.interpares.org/book/interpares_book_d_part1.pdf.

[13] Evelyn Peters McLellan (2006), "InterPARES 2 Project - General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_english.pdf (English); http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_french.pdf (French).

Case Study Data

The case study data were incorporated into the research of the Policy Cross-domain in three stages. The first stage was a review of the responses contained in each of the case study reports to the four questions below:

20. To what extent do policies, procedures and standards currently control record creation, maintenance, preservation and use in the context of the creator's activity? Do these policies, procedures and standards need to be modified or augmented?

21. What legal, moral (e.g., control over artistic expression) or ethical obligations, concerns or issues exist regarding the creation, maintenance, preservation and use of the records in the context of the creator's activity?

22. What descriptive or other metadata schema or standards are currently being used in the creation, maintenance, use and preservation of the recordkeeping system or environment being studied?

23. What is the source of these descriptive or other metadata schema or standards (institutional convention, professional body, international standard, individual practice, etc.)?

The second stage involved another pass through the case studies with targeted explanatory prompts on intellectual property and privacy issues. It was executed by a small team of research assistants in late 2004, with the benefit of the policy studies then completed or nearing completion. The third stage is the composition of this report. This has involved a final review of data gathered as well as providing case study leaders with the opportunity to comment on the conclusions reached.

Except for this final stage, matching the fragmentary policy data coming from the case studies with those from the higher level studies was problematic. This can be partly attributed to the difficulty of interdisciplinary exchange between the perspectives of archival science and political science.

Policy Themes

The record creation environments that emerged from the case studies and the regulations for record creation, maintenance and preservation that emerged from the policy studies show patchy-to-nonexistent degrees of maturity. In terms of a comprehensive framework, few organizations and legislative jurisdictions show that they can deal adequately with the digital challenge, particularly where management requirements at any time in the records lifecycle involve complex multi-component records as appeared in many of the case studies. The main exception to this grim picture was, unsurprisingly, the sphere of evidence law: across a wide range of jurisdictions, legislation related to the use of records as evidence in a court of law shows a considerable congruity with the findings of the Authenticity Task Force of InterPARES 1.

Answers to the original research questions of the Policy Cross-domain follow in the next section. The four statements reflect the principle policy themes that emerged from the research done by the Policy Team, by other groups within the Project and from the case study reports.

Theme 1: An inclusive policy infrastructure for recordkeeping is required to support the activities of a society heavily reliant on information technology.[14]

A principal finding of InterPARES 2 is that preservation of records emerging from interactive, dynamic and experiential environments requires an inclusive policy infrastructure beyond the principles expressed by the InterPARES 1 Strategy Task Force. Concerns of intellectual property, privacy and security pre-exist the digital recordkeeping environment, but in relation to a minority of records. These concerns are now far more prevalent. The Policy Team has presented the necessary elements of the top level of such an infrastructure in the Framework of Principles for the Development of Policies, Strategies and Standards for the Long-Term Preservation of Digital Records. These are capable of implementation at a variety of levels of governance.

In the current networked/inter-connected environment, the following concerns become central because of the increasing transfer of information across organizational boundaries. While the first three concerns are external to records and traditionally provide the basis for archival preservation, the intellectual property rights concern applies to the record both externally (to its context) as well as internally (to its content); the last two concerns are internal to the record.

- Relationship to business process
- Relationship to specific transaction
- Relationship to creator
- Relationship to intellectual property rights (context and content)
- Relationship to privacy (content)
- Relationship to security (content)

Acceptance of a new conceptual understanding of the nature of the record that is extensible to these new environments and its use in tandem with the related policy principles will encourage the commonality of approach required to turn InterPARES' theoretical outputs into a robust foundation for formal standards development and policy creation at organizational, sectoral, national and international levels. Like any policy principles and any juridical instruments designed to support them, standards need to be facilitative and not specific to any particular technology to be useful.

A research project such as InterPARES has to define best or even ideal practice based on clearly articulated theoretical principles. Standardization in the national or international arena tends to focus on either setting acceptable baseline requirements or formalizing commonly accepted "best" practice as a norm. The proliferation of computing has tended to nudge standard-setting towards the second whereas arguably it ought to be confined to the first. This is particularly true in the area of promoting interoperability—across time and space—between digital systems, which is vital to support information transfers and consequently the archival process. Aside from the study of file formats already cited as spun off into Domain 3, the Policy Cross-domain has not directly addressed the development of standards. However, other research units within the Project have observed and collaborated with standards-setting initiatives.[15]

[14] In the InterPARES 2 Terminology Database, "recordkeeping" is defined as "The whole of the principles, policies, rules and strategies employed by the creator that establishes and maintains administrative, intellectual and physical control on its records.," and "recordkeeping system" is defined as "A set of rules governing the storage, use, maintenance and disposition of records and/or information about records, and the tools and mechanisms used to implement these rules" (http://www.interpares.org/ip2/ip2_terminology_db2.cfm). These concepts are reflected in the InterPARES 2 Chain of Preservation (COP) Model (see the Modeling Cross-domain Task Force Report).

[15] Throughout the duration of InterPARES 2, there have been significant developments in the digital longevity standards arena, particularly open standards. ISO 19005 (see International Organization for Standardization, ISO 19005-1:2005 - Document

Theme 2: An expanded and more detailed definition of record is necessary.

InterPARES 2 findings recommend preservation of all documents the creator treats as records; that is, all documents that the creator relies upon in the usual and ordinary course of affairs, associates with other records and refers to as the records of its affairs. This is more consistent with the inclusive definition of the term “record” used in statutes. It is the creator’s judgement of what constitutes the record to be kept for action and reference, and the preserver has then to assess the feasibility of preserving it over the long term.

InterPARES 2 findings also point to a new category of records: potential records. Records have traditionally been identified as such retrospectively; that is, after having been completed and issued with a fixed form and stable content; but, with dynamic systems, there is the possibility of identifying “prospective” records. The digital objects that clearly manifest themselves as records since the moment they are created fulfil the traditional, memorial function of records to bear witness to or remember an action in which they participated or of which they were the residue. Rather than witnessing the past, prospective records guide the future through a set of instructions or actions to be carried out.[16] As such, prospective records may not be considered records when their process of development begins, but, since their content can be fixed and their documentary form and functionalities described to make it possible to re-create them in the future, they could become records. Establishing policies to manage recordkeeping for digital objects that are prospective records and *may* become records appears to fall into the context of guides, manuals and other directive or procedural documents.

Theme 3: Business processes are divided between many systems.

Deployment of dynamic, interactive and experiential systems to capture, handle and manage data is at present not always undertaken with due consideration of the various roles of records (as memorials; i.e., for reference, or as directions; i.e., instructions for future activities), nor of their

management—Electronic document file format for long-term preservation—Part 1: Use of PDF 1.4 (PDF/A-1)) is a file format specification derived from the PDF Reference, Third Edition, version 1.4 of Adobe Systems Incorporated’s commercial software *Acrobat* (a matrix image format with some textual support capability). This is an encouraging example of a proprietary software format specification becoming an openly available specification once the owners of the intellectual property have replaced it with another format for their main revenue-generating markets. In this case, the intellectual property is to be managed by ISO for fifty years. In late 2006, Microsoft Corporation announced that future versions of its Microsoft Office System software, beginning with the 2007 version, will support saving documents in its XML-encoded “Office Open XML” (abbreviated as OOXML), which is a file format specification created by Microsoft for the storage of digital documents. The format was standardized by Ecma (*European Computer Manufacturers Association)* International as Ecma 376 in December 2006, which has since been submitted for adoption under the ISO/IEC JTC 1 process. An important distinction should be drawn from the archival perspective between widely adopted “industry” standards and those that are genuinely open: dependencies on the current computing environment may exist for both record content and metadata at encoding/syntactic, computer file, application and database levels as well as computer hardware. There has been formal collaboration between researchers in the Description Cross-domain developing the Metadata and Archival Description Registry and Analysis System (MADRAS) (see http://www.gseis.ucla.edu/us-interpares/madras/guidelines.php) and the working group within ISO Technical Committee 46/Sub-Committee 11 drafting the third part of ISO 23081 - Information and documentation—Records management processes—Metadata for records. MADRAS is a tool that is designed to increase the visibility of recordkeeping and archival metadata schemas and to facilitate the comparison of schemas against established requirements. Similarly, in the fall of 2006, the Project made a submission to the revision of ISO 14721, the *Open Archives Information System Reference Model*. Within the case studies, the most comprehensive and policy-driven observance of record creation standards was found to be in the science focus and especially case studies 06 and 19. Many of the government focus case study reports mention standards in response to direct case study questions, but they are either substantially irrelevant to recordkeeping and preservation requirements or not actually implemented (e.g., ISO 15836:2003 - *Information and documentation—The Dublin Core metadata element set* was frequently cited).

[16] Luciana Duranti and Kenneth Thibodeau (2006), “The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES,” *Archival Science* 6(1): 13–68 (Note: a reprint of this article is provided in Appendix 2).

sometimes distributed nature. Systems may be distributed across an organization, which may itself be distributed (e.g., multi-national organizations that cross national boundaries). Both of these scenarios are the norm in collaborative e-Science projects such as the Cybercartographic Atlas of Antarctica examined in case study 06.

Dynamic, interactive and experiential systems may also be deployed to achieve objectives that are not compatible with those of recordkeeping (e.g., providing a “window” on existing data at a particular point in time, as in the VanMap case study). Documents that may satisfy recordkeeping requirements can be instantiated at multiple points across modern systems. The Revenue On-Line Service (ROS), examined in case study 20, was essentially a conduit, enabling controlled input of data directly by citizens rather than by government staff working from paper forms mailed in by citizens. The digital objects of the ROS are records meant to establish and normalize a citizen’s relationship with the revenue agency. That is, the business of citizens paying taxes was broken out into at least two systems: the ROS, which managed the relationship of the citizen with the revenue agency, and the mainframe computers, which actually assessed the taxes. In addition, documentary elements that convey the semantic of a record (metadata schemas, for instance), may exist in systems as digital objects separate from the record. In case study 19, an engineering experiment that used Web Ontology Language (OWL), an extension of XML that allows representation of semantics within metadata schemas to formulate a new logical preservation format for complex CAD records, metadata elements were stored in a segment of a pilot preservation system located on the opposite end of a national network shared by the experiment partners.[17]

The sub-division of a business process between systems, some of which may (a) be dynamic, interactive, or experiential and (b) exist across organizational boundaries, suggests a need for policy direction as comprehensive as the systems and business processes at hand. Records identified in one system must be considered along with records related to the same business process created by other system(s) to ensure the most effective management, disposition and preservation of records takes place. Policy should ensure that: (1) the identification of documentary entities, including but not limited to records/metadata/linkages, etc.,[18] is undertaken at the system design phase, (2) appropriate functions are incorporated to manage and preserve the entities identified at the outset of system development and (3) the process and outcomes of these activities are reviewed regularly as part of system operations.

Theme 4: Preservation policies are inadequate or absent.

The digital objects considered to be records by their creators may not be preservable because of poorly considered use of digital signatures, for example, or may otherwise not be fit for functioning as retrospective or prospective records. The preservation activities examined by the case studies were mostly directed at keeping the data, not the record. For example, back-up and disaster recovery routines were found to be widespread, but the ability to restore the records except to an identical system (interoperability across time) was rarely addressed. In science Focus

---

[17] See Kenneth Hawkins (2006), “InterPARES 2 Project - Case Study 19 Diplomatic Analysis: Preservation and Authentication of Electronic Engineering and Manufacturing Records.” Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs19_diplomatic_analysis.pdf; and Kenneth Hawkins (2006), “InterPARES 2 Project - Case Study 19 Final Report: Preservation and Authentication of Electronic Engineering and Manufacturing Records,” 14, 18. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs19_final_report.pdf.

[18] Because semantic value can be derived from an understanding of how documentary entities relate to one another (for example, a registry to a series of records and the records themselves), additional entities of interest might include data and system models, domain-specific taxonomies and enterprise architecture models and specifications.

case study 06, the Cybercartographic Atlas of Antarctica, steps taken to ensure interoperability across systems performed many of the same purposes as preservation. Similarly, production of exhibits that are accessed via the Web may include elements of a preservation policy, albeit not a long-term one, because the produced objects are considered to be maintained accurate and authentic across space (i.e., from one system to another). Consideration of what constitutes a preservation policy needs to be inclusive, as shown by the Chain of Preservation Model.[19]

Addressing the Research Questions

The research statement of the Policy Cross-domain contains the following questions, to which the report provides a combined response to avoid repetition.

- To what extent do policies, procedures and standards currently control record creation, maintenance, preservation and use in each focus area? Do these policies, procedures and standards need to be modified or augmented?
- Can an intellectual framework or frameworks be developed to facilitate the translation of policies, procedures and standards into different national environments, sectors and domains?
- How can enhanced control over and standardization of record creation, maintenance, preservation, access and use be balanced against cultural and juridical differences and perspectives on issues such as freedom of expression, moral rights, privacy and national security?
- What legal or moral obligations exist regarding the creation, maintenance, preservation and use of the records of artistic and scientific activities?
- What principles should guide the formulation of policies, strategies and standards related to the creation of reliable, accurate and authentic records in the digital environments under investigation?
- What principles should guide the formulation of policies, strategies and standards related to the appraisal of those records?
- What principles should guide the formulation of policies, strategies and standards related to the long-term preservation of those records?
- What should be the criteria for developing national policies, strategies and standards?
- What should be the criteria for developing organizational policies, strategies and standards?

Recordkeeping and the current policy environment

The degree to which policies, procedures and standards control record creation, maintenance, preservation and use in the three focus areas examined varies from none at all to partial. It appears that the two key factors affecting the response to this research question are the nature of the organization and the phase or stage of the records lifecycle (i.e., record creation, maintenance, etc.), being considered.

Where organizational culture is conducive to the development of policies and procedures and the adoption of standards, there are controlled aspects of record creation, maintenance, preservation and use. Of the environments studied, government and some scientific organizations

[19] See the Modeling Cross-domain Task Force Report.

developed or adopted policies and procedures, while individual artists or small, temporary partnerships did not. In these latter organizations, the extent of control on recordkeeping through policies, procedures and the adoption of standards is effectively nil.

In half of the arts focus case studies, it was found that there was no consideration given to record maintenance and preservation. Among the remaining case studies, where there was some consideration given, there was no common motivating factor. In some instances, record maintenance and preservation were motivated by legal reasons, usually pertaining to the protection of intellectual property. Publicity and future performances were two other motives for record maintenance and preservation, although this did not necessarily extend to the development of policies or adoption of standards.[20]

Among the science focus case studies, it was noted that the development of rules and procedures around record creation and maintenance is driven by the immediate and foreseeable requirements of each scientific activity. It was also found that while sophisticated technologies were often adopted, those used to maintain and access the records tended to be rudimentary (e.g., Microsoft Windows or other proprietary software tools). A process to determine how long to keep project data was consistently in place. Procedures supporting retention included duplication and migration. While there was no consistent approach or procedure to achieve the determined retention requirement, enabling others to access the data on different systems (i.e., interoperability), was frequently a guiding consideration.

Of the three categories of case studies, those involving government organizations consistently had the most comprehensive recordkeeping policies and procedures in terms of all phases of the lifecycle. In most cases, organizations had an existing, formal relationship with an archives or other unit responsible for record preservation. As with the science focus case studies, however, it was found that maintenance and preservation processes were data- or system-oriented and not necessarily linked to the organization's specific recordkeeping requirements.

Another factor to consider in relation to where an expressed policy exists or standards are adopted is the phase of the records lifecycle. The case studies showed that organizations may adhere to policies, procedures and standards in one stage of the lifecycle (e.g., record creation), but not in others. The InterPARES concept of records preservation is comprehensive and includes all activities that affect the record since its creation. None of the organizations involved with the case studies displayed such a comprehensive approach.[21]

A study of records-related legislation concluded that laws in the jurisdictions studied are very inclusive in their definitions of records, in contrast with the much more specific archival definition adopted by the InterPARES 2 Project:

> A document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference.[22]

Inclusive and inconsistent definitions of record undermine not only an organization's ability to develop the policies and procedures it needs, but also decisions to adopt existing or proposed standards. They also compromise an organization's ability to correctly interpret precedents set in court decisions regarding records.[23]

[20] See the Domain 1 Task Force Report.
[21] The conclusion that preservation of digital records must begin at the creation stage has been reached by most, perhaps all, research in this field. This conclusion is thoroughly developed in InterPARES 1 and in that Project's strategic principles as follows: "…preservation of authentic electronic records is a continuous process that begins with the process of records creation..." (Duranti et al., "Strategy Task Force Report," op. cit., 4).
[22] Definition for "record" from the InterPARES 2 Terminology Database, op. cit.
[23] Examples exist where organizations have been sued for large amounts while other, similar organizations continue to ignore the risk.

The study of records-related legislation concluded that while all phases of the records lifecycle are addressed in the laws or directives of the jurisdictions studied, they are not addressed comprehensively within any single law, nor overall within the body of legislation examined within any one jurisdiction.[24] For example, land transactions are a highly regulated business activity. In Alsace-Moselle (France), case study 18, information technologies were adopted to carry out this activity. The system developed is very effective for the short to medium term, but presents unresolved long-term maintenance and preservation issues, especially with regard to maintaining the authentication function of the judge's digital signature over the long term. The same study also concluded that the records lifecycle phases most commonly addressed within legislation are those of creation and disposition.

Statutory recordkeeping requirements provide a strong impetus for organizations governed by them, but it cannot be expected that legislation will consistently and comprehensively address all phases of the records lifecycle. In the absence of a comprehensive guidance or direction from law, organizations may be willing to adopt general standards, such as the records management standard,[25] to help them effectively maintain and preserve their records.

Balancing cultural differences against a common approach

The widespread adoption of new technology in the three environments examined by InterPARES 2 does not appear to have necessarily or fundamentally changed the long-standing processes in those environments. The Domain 1 Task Force Report on record creation concludes that the processes occurring in the creation of records today are recognizable in those used in the pre-digital environment.

> What is being witnessed in the arts focus case studies is the *continuation* of the artistic tradition in the digital environment. The processes are largely the same, based on the long-established artistic principles of each field… [for most of] the science focus case studies, document creation takes place in a much more formalized and controlled environment, with pre-determined processes including the collection, analysis and preservation or communication of data…Most of the case studies in the government focus deal with a traditional activity being carried out in a new way. Therefore, the process of document creation is largely the same as for the traditional environment; it is simply transposed into the digital environment with the possible addition of certain steps in the process to take the technology into account.[26]

If the processes are not changing at a fundamental level, then existing policies, strategies and standards may not need to be completely changed, but simply revised and extended. In the pre-digital environment, where records and physical media are inseparable, the point at which a record is created is well established (e.g., a film is made or a letter is written). InterPARES 1 concluded that in the digital environment the medium is no longer an essential part of a record. As a consequence, preservation must be directed to preserving the ability to reproduce digital records, moving them, as needed, from one medium to another. Therefore, record creation procedures and standards must set out when a record has been created as well as identify the intellectual and digital components comprising the record and their relationships to each other. For records

[24] Suderman et al., "Archives Legislation Study Report," op. cit., 24.
[25] See International Organization for Standardization, ISO 15489-1:2001 - Information and documentation—Records management—Part 1: General.
[26] Domain 1 Task Force Report, 89, 91, 92 (emphasis in original).

created or existing in dynamic systems, procedures must outline how those components are determined and set out the acceptable range of variations on their relationships to each other.[27]

New principles that guide policies, procedures and standards on the identification and modification of created records are also required. Digital technologies have dramatically increased the opportunity to integrate record types formerly distinguished by their media (e.g., audio and text). Besides the well-known complications linked to maintaining and/or preserving the differently formatted digital components of which records are comprised, this capacity requires a significantly enhanced management of intellectual property rights existing within the records. These may include database rights, copyrights and patents. Similarly, the emergence of access and privacy legislation requires a more comprehensive management of record content than existed before. Laws governing personal information emphasize accuracy and enable a person identified within the record to request that information within a record be corrected. Rights inimical to the preservation of the records by a preserver may subsist at the record component level. Whereas previously the preserver could manage these issues by considering "sunset" periods for which entire records might be withheld until the rights had expired, they now need addressing in policies and corresponding rules applying from creation. Records creators must also have clear procedures in place for how those rules are implemented. These procedures must be explicitly understood by the individuals responsible or be built into the design of systems that maintain the records.

Digital technology has also dramatically enhanced the means to transmit information. In both the arts focus and science focus case studies, this was found to be a welcome characteristic. For the artist, the ease of transmission can dramatically increase the potential audience for a created work. For scientists, greater access to more data supports more effective research. The scientific community is motivated by "the desire and possibility of translating the collected data into a neutral or open source format."[28] By contrast, security concerns predominate in the governmental environment, where record transmission is emphasizing the need for security metadata and technologies to support legal non-repudiation by participants in the record creation process, as well as standards for secure storage technologies, such as encryption, secure digital signatures and biometrics. Governments also exchange information, of course. For records to be correctly accessed across space, explicit policies are required not only so that the receiver of the transmitted record can accurately reproduce it, but also so that records sent do not contravene requirements for security, privacy protection and intellectual property in either jurisdiction. There are three generic approaches to achieving this greater communication with appropriate safeguards: (1) harmonizing juridical frameworks, (2) implementing effective exemptions for the purposes of archival preservation and (3) ensuring comprehensive rights metadata accompany the record.

InterPARES 1 emphasized the importance to the preserver of assessing the feasibility of preservation during the records appraisal process. Feasibility assessment policies and procedures need to be guided by the technological requirements of the records as they relate to the capabilities of the preserver's preservation system. They must also take into account what residual rights or obligations—privacy, intellectual property, security, etc.—will have to be managed or administered by the preserver. In this way feasibility operates both specifically (i.e., in relation to an identified body of records), and generally, in that the preserver must develop or modify acquisition policies so that they are consistent with the capabilities of the preservation system.

[27] See Duranti and Thibodeau, "The Concept of Record," op. cit.
[28] Domain 1 Task Force Report, 91.

The preserver who is maintaining authentic copies of created digital records must, in effect, be guided by the same concerns as the creator. That is, if the creator had to observe requirements of privacy, intellectual property and security while maintaining the records, the preserver must also observe those requirements within the preservation environment, unless explicitly exempted. The foremost principle that must guide the long-term preservation of digital records was established in InterPARES 1, which is to ensure that through preservation processes records remain authentic copies of the creator's records.

The literature reviewed in the annotated bibliography of intellectual property rights points to the issue of enhanced control over access to digital content in the service of commerce as a key one at play in the formulation of international treaties, national legislation, case law and policy debates. Ironically, the very same technical and legislative features that enable economic protection for rights owners and enhance the immediate access to records and information for consumers also make more restrictive the future uses of the content and ultimately impede the ability to preserve these for their "second noncommercial life."[29] The emergence of access and redistribution control technologies (also known as digital rights management or DRM) and attendant debates about these technologies and the challenges they introduce for preservation demonstrate well the precarious balance struck between enhanced control over access, communities' expectations and juridical perspectives on use.[30]

Reflecting legal and moral obligations in policy

Activities of records creators and records preservers are subject to legal and moral obligations as well as community expectations. Records preservers are "downstream" recipients of evidence of the activities of records creators of yesterday and today. At the same time, records preservers are not only recipients of records but also records creators in their own right. Where in the past records preservers, especially archives, managed the transfer of physical and intellectual property rights in analogue records, the picture is quite different today. As records creators use software to create and/or apply rights management technologies to wrap or otherwise protect intellectual assets, they introduce a whole new layer for preservation management. Because this additional layer is itself subject to intellectual property rights and protections, the process of preservation takes on additional tasks and risks.

The emergence of access and redistribution control technologies comes at a time when moral rights are being trumped by commercial rights, and privacy rights are being overwritten by assertions of national security. Successive changes to national laws, international trade agreements and business models render the already considerable challenge of preserving digital records far more complex than simply overcoming issues of technological obsolescence.

Beyond ensuring that preserved digital records remain authentic copies of the creator's records, preservation should be seen and undertaken as a process that is compatible with the purpose for which the preserved records were created. If compatibility of purpose cannot be established, preservers may require specific or general exemptions from liability under intellectual property requirements, including moral rights, and privacy requirements. For

[29] "Second noncommercial life" has been elaborated by legal scholar Lawrence Lessig as the period that begins when the copyright term expires and content becomes subject to re-use; see Lessig's *Free Culture* (The Penguin Press, 2004), and the Editorial, "The Coming of Copyright Perpetuity," *New York Times*, January 16, 2003, p. A28.

[30] For instance, see the Canadian Internet Policy Clinic (University of Ottawa) policy debates, available at http://www.cippic.ca/en/faqs-resources/digital-rights-management/, and "Digital Rights Issues" in the American Library Association, Washington Policy Office, available at http://www.ala.org/ala/washoff/WOissues/copyrightb/digitalrights/DRMissues.pdf.

example, anonymization of records containing personal information compromises the integrity of the created record. Preservation activities may result in changes to the records at the bit level, but not at a functional level. Such activities would contravene a rigid application of intellectual property rights.

The records created in some of the arts focus case studies were insufficiently well-defined for a preserver to demonstrate the authenticity of reproduced copies of the records. In *Obsessed Again...*, case study 13, a reproduction of the work was deemed not to be authentic by the creator. This suggests the importance of the principle that preservers must interact with creators from the outset to preserve authentic copies of records. Where that relationship does not exist, the preserver's procedures and standards must set out the extent of the authenticity of the reproductions of preserved copies.

Principles for appraisal and preservation

The combination of rapid technological change and the need to manage rights subsisting within the content or components of the records poses significant challenges to the long-term preservation of digital records. The development and adoption of common standards and stronger procedural controls for recordkeeping cannot by itself enable long-term preservation. A clear and ongoing relationship between creator and preserver is also necessary. Each aspect sustains the other. Standards and procedural controls inform the selection of record creation and maintenance technologies by the creator. The development of guidelines and procedures and the adoption of standards would comprise the principal aspects of the preserver's participation at the record creation phase. Established procedures and standards also help the preserver to develop and operate a preservation system and, much more importantly, demonstrate the authenticity of records maintained in that system. The European Union

> consider[s] standardization "an integral part of their policies to carry out 'better regulation', to increase competitiveness of enterprises and to remove barriers to trade at international level."[91] The directives on *Data Protection*, *Electronic Signature*, *e-Invoicing* and the regulatory framework for electronic communications networks and services (which consists of five additional directives) are a set of new legislation (categorized as Information Society legislation). They are issued under the aegis of the European Standards Organizations with the purpose of establishing a "legal framework to ensure the free movement of information society services between Member States."[92] [31]

Standards and procedural controls are static in relation to the creativity of record users and deployment of new technologies and systems and so, by themselves, cannot accommodate national and cultural differences.[32] A sustained relationship between records creator and preserver is a means by which the creator can communicate innovative uses or procedural variations to the preserver. Such a relationship also informs the preserver of ethical behaviour of a community (e.g., that in the scientific research community, research data are to be shared as broadly as possible, but not until those who have prepared or gathered the data have had a reasonable opportunity to publish their findings).

[31] Suderman et al., "Archives Legislation Study Report," op. cit., 30. Note: footnote references in the quote are from the original text, and are not reproduced here.

[32] Case studies used in both phases of InterPARES were from many different jurisdictions. For specific jurisdictional studies on specific issues, such as authenticity, see the Policy Cross-domain studies summarized above in the section entitled "Research Methodology."

The extent to which public, sectoral and organizational policy affects the participants in various record creating activities varies according to the legal, ethical and moral dimensions of their relationships with their correspondents. Thus, there are fewer legal and moral obligations affecting recordkeeping in the artistic and scientific environments in comparison with obligations existing in the governmental environment. Those in the artistic environment might be summarized as relating to intellectual property, while obligations in the scientific environment would centre more on accuracy and accessibility of research data.[33]

Among the arts focus case studies, some creators simply were unconcerned with the long-term risk of loss of their digital records. As noted above, even where the creating organizations were concerned with recordkeeping, obligations related to protecting or acknowledging intellectual property or meeting financial accountability requirements to a granting body.

Legal and moral obligations for recordkeeping may be increasing within the scientific community, driven primarily by policies of funding organizations. As a knowledge-based community, it is in its own interest to ensure that research data be maintained for future use. Some scientific communities have long-established recordkeeping standards (e.g., metadata standards for geospatial data), to which the community expects researchers to adhere. Likewise there is widespread use of creative commons licenses to support general access to and use of scientific research data while simultaneously establishing the ownership rights of the creator.

Recent legislation governing personal information has extended obligations in this regard beyond the governmental sector and highly regulated private sector activities, such as banking. This legislation imposes some additional obligations on any organization collecting and using personal information. Research communities already have rules in place for the ethical collection, use and maintenance of research data involving human subjects, so it is safe to say that legal obligations are increasing for records containing personal information in the scientific environment. It is unclear how this new legislation will affect records in the artistic environment.

Intellectual property laws pose a dilemma for records preservers where such rights subsisting in the preserver's custody may be aggressively protected. The moral obligation not to change the creation of an artist, for example, may make long-term preservation impossible if the created record relies on short-lived technological components. In some jurisdictions, laws have exempted specific preservation institutions, usually national archives, from liability arising from copyright (e.g., Library and Archives Canada is explicitly allowed to "crawl" and capture Canadian Web content). In terms of personal information, preserving organizations or their clients may need to prove to an external authority that their use of records containing personal information is compatible with the purpose for which it was created.

Importance of a common basis for national policies

National policies, strategies and standards should be guided by common approaches and common purposes for all phases of recordkeeping. The directives developed in the European Union (EU) are an example of how common criteria can be set for specific implementation within each member nation. Criteria developed in this way need to be reviewed or measured against technological and economic limitations. While the general principles of the EU's e-Signature Directive are being implemented by member states "despite recognized limitations in the technology supporting e-signatures [it is observed that] 'there is currently no market demand

[33] The link to the recordkeeping practices of individuals is in their relationships with organizations, and InterPARES guidelines for their own recordkeeping and preservation are referred to in the *Policy Framework* (see Appendix 19).

for qualified certificates and related services.'"[34] The limitations of technology are reflected in the caution shown in the EU's e-Signature Directive, which "explicitly excludes certain categories of contracts."[35]

The relationships between records creators and preservers must be acknowledged and supported by national policies, strategies and standards. This will require rules for recognizing professionals and organizations that preserve digital records throughout the society, not just in terms of national institutions. Such rules will need to address the obligations preservers must meet; that is, make explicit the characteristics of a trusted custodian, in connection with the rights subsisting within the records being preserved and the transient nature of technologies used for recordkeeping, particularly those for record creation. National policies and standards must also be flexible enough to accommodate the norms of specialized communities, such as the creative arts and scientific research, which may themselves not be particularly bounded by national borders.

While it is recognized that national policies and standards need not be comprehensively addressed in legislation, it is important that legislation be developed within the most comprehensive information strategy possible. Establishing policies for which consistent implementation is impossible, as in the case of the European Union's e-Signature Directive, jeopardizes the rights of all. Without authentic, reliable and accurate records and rules about their use and transmission, rights such as those pertaining to privacy or intellectual property of citizens may be violated. Without record creation, maintenance and preservation policies in place, the state may itself participate in the violation of those rights.

Criteria for organizational policies

Organizational policies, strategies and standards for recordkeeping must obviously meet legislated requirements. To the greatest degree possible, organizations or communities of practice must codify how these requirements will be met. Where legal obligations subsist within records (e.g., privacy or intellectual property) and where these would be contravened by normal maintenance or preservation activities, organizational recordkeeping policies should incorporate a risk assessment component. Organizations may protect themselves at least to some degree by working collectively with similar organizations to establish common practices. Such an approach will necessarily involve consideration of all phases of recordkeeping and all organizations within the community participating in any of those phases.

Another criterion for organizational recordkeeping policies, strategies and standards is the explicit consideration of long-term preservation requirements. This is essential not only to determine what those are, but also to determine whether preservation is even possible or desirable within the organization or whether an external preserver must be identified. In the latter case, the presence of long-term preservation requirements will form the basis of the relationship between the creating and preserving organizations.

Two sets of guidelines were produced by InterPARES 2 to assist individuals and organizations with establishing recordkeeping policies, strategies and standards. The first set of these is entitled *Creator Guidelines—Making and Maintaining Digital Materials: Guidelines for Individuals*,[36] and is intended to help individuals or small organizations who are making and maintaining digital materials, including records. As the case studies revealed, "the technology

[34] Suderman et al., "Archives Legislation Study Report," op. cit., 31.
[35] Ibid., 31–32.
[36] See Appendix 20.

used by innovators and early adopters, regardless of the focus area in which they belonged, was proprietary and frequently customized," and that "[i]n many cases, the point of the work of these types of creators is to explore, test and push the limits of the available technology, be it hardware or software."[37] These guidelines are intended to inform creators who may not consider or be aware of digital record creation and maintenance concerns. In particular, evidence of authorship, with implications for preservation of intellectual property rights, is at risk even if technological obsolescence issues are addressed.

The second set of guidelines, *Preserver Guidelines—Preserving Digital Records: Guidelines for Organizations*,[38] provides more procedural guidance for any organization charged with providing preservation services (i.e., where preservation considerations are central to the organization). These guidelines are not specific to large, established archival organizations. They support the development of preservation procedures and systems that can maintain the accuracy and authenticity of the preserved records and are conceptually linked to the components described in the InterPARES 2 Chain of Preservation model.

Toward an Intellectual Framework for Policy Development

It is clear from the foregoing that not only can more comprehensive policies, procedures and standards be developed, but also that they are needed. Several of the products developed by InterPARES 2 contribute comprehensive guidance for all aspects of digital recordkeeping. In terms of policy, the main product of the Policy Cross-domain is the already mentioned Framework of Principles, comprising two complementary sets of principles for the creation and preservation of digital records. These principles are introduced and detailed in Appendix 19. The Framework provides the scope for developing a consistent and comprehensive policy environment in different jurisdictions, sectors and organizations. It may also help with the assessment of standards, existing and contemplated, and with the development of new standards, in terms of their applicability and utility for all aspects of recordkeeping.

The Framework extends the strategic principles established in InterPARES 1 in three key ways.[39] Firstly and most importantly, it sets out principles for record creation from the creator's point of view. Secondly, because of the Framework's dual viewpoints (creator and preserver), it structures the relationship between a records creator and preserver. This relationship is seen as one of gradually transferring responsibility from the creator to the preserver.[40] While the idea of a shifting responsibility is not new, the principles clarify the dynamics of a seamless transfer. Thirdly, the Framework considers records emerging from three different environments (i.e., the arts, the sciences and government/administration) and which exist in dynamic, experiential or interactive systems. The scope of its application in terms of organization and system is thus much more inclusive than the strategic principles from InterPARES 1, which were based on governmental organizations where the most comprehensive recordkeeping policy environments existed.

[37] Domain 3 Task Force Report, 178–179.
[38] See Appendix 21.
[39] See Duranti et al., "Strategy Task Force Report," op. cit.
[40] The theory of movable responsibility was developed decades ago in the United Kingdom by Felix Hull. It recognized that the records manager and archivist worked together throughout the records lifecycle but with the responsibilities of the former gradually diminishing as those of the latter grew.

PART EIGHT

TERMINOLOGICAL INSTRUMENTS

Terminology Cross-domain Task Force Report

by

Joseph T. Tennis, University of Washington
Randy Preston, The University of British Columbia

Introduction

When research is carried out by a multidisciplinary and multicultural team that spans more than one dozen fields of inquiry and twenty countries, the precision and consistency of the terminology used in the course of the research is vital to the success of the Project. Several terms that are key to this Project refer to different concepts in many of the disciplinary and/or cultural environments involved, while similar concepts are expressed by different terms. The term "record" is but one example. As used in the context of the InterPARES 2 research, a record is defined as "a document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference."[1] This definition stands in stark contrast to that used by most of the other disciplines involved in the Project. For example, in Computer Sciences, a record is often defined as "an ordered set of fields, usually stored contiguously" or "a grouping of interrelated data elements forming the basic unit of a file," while in the Arts a record refers to "any electronic, photographic or mechanical recording of music, singing, dialogue, sound effects or visual events, including CDs, DVDs, audio tapes, films, videos and the like."[2]

The InterPARES 2 Terminology Cross-domain Task Force was responsible for researching and defining all terms proposed for official use by each research unit within the Project and accepting or rejecting them on the basis of clarity, consistency with the other adopted terms and validity in the various disciplinary and cultural contexts. To this end, the Task Force developed a Terminology Database composed of three terminological instruments designed to be of service to each of the other research units within the Project and, by extension, to Archival Science. Over the course of the five-years of the Project, the terminology team collected words, definitions and phrases from extant documents, research tools and models, and through direct researcher submissions and discussions. From these raw materials, the team developed a systematic and pragmatic way of establishing a coherent view on the concepts involved in dynamic, experiential, and interactive records and systems in the arts, sciences and government.

Research Team

Team composition

Like each of the Project's other cross-domains, the Terminology Cross-domain comprised researchers from a mixture of academic, archival and cultural heritage institutions, assisted by graduate research assistants from the universities of British Columbia and California, Los Angeles. During the first three and a half years of the Project, the Cross-domain was chaired by Jonathan Furner of the University of California, Los Angeles. In mid-2005, after Furner stepped down as chair, acting chairmanship was passed to Luciana Duranti before a new chair, Joe Tennis, was appointed in September 2005.

The following is a complete list of researchers and research assistants who participated in the Terminology Cross-domain Task Force at some point during the Project.

---

[1] InterPARES 2 Terminology Database. Available at http://www.interpares.org/ip2/ip2_terminology_db.cfm.

[2] Ibid.

Chairs:

Jonathan Furner	Jan 2001 - June 2005
Luciana Duranti	June - Sept 2005
Joe Tennis	Sept 2005 - Dec 2006

Researchers:

Barbara Craig	University of Toronto, Canada
Luciana Duranti	The University of British Columbia, Canada
Philip Eppard	University of Albany, State University of New York, USA
Jonathan Furner	University of California, Los Angeles, USA
Ian Lancashire	University of Toronto, Canada
Richard Pearse-Moses	Arizona State Library, USA
John Roeder	The University of British Columbia, Canada
Joe Tennis	University of Washington, USA
James Turner	Université de Montréal, Canada

Research Assistants:

David Boudinot	The University of British Columbia, Canada
Natalie Catto	The University of British Columbia, Canada
Naomi Cull	The University of Toronto, Canada
Kimberly Davison	The University of British Columbia, Canada
Shanna Fraser	The University of British Columbia, Canada
Jessica Glidewell	The University of British Columbia, Canada
Nadine Hafner	The University of British Columbia, Canada
Peggy Heger	The University of British Columbia, Canada
Eleanor Kleiber	The University of British Columbia, Canada
Tracey Krause	The University of British Columbia, Canada
Karen Langley	The University of British Columbia, Canada
Yvonne Loiselle	The University of British Columbia, Canada
Katherine Miller	The University of British Columbia, Canada
Emily O'Neill	The University of British Columbia, Canada
Carolyn Petrie	The University of British Columbia, Canada
Randy Preston	The University of British Columbia, Canada
Corinne Rogers	The University of British Columbia, Canada
Nadav Rouche	University of California, Los Angeles, USA
Anthea Seles	The University of British Columbia, Canada
Brian Trembath	The University of British Columbia, Canada
Melanie Wallace	The University of British Columbia, Canada
Catherine Yasui	The University of British Columbia, Canada

Research mandate

The InterPARES Project set out the following responsibility for the Terminology Research Team:

> The Terminology Research Team will control the use of terms in all areas of the research. This research team will establish formal procedures for the proposal and adoption of specific terms, and meet in conjunction with International Team workshops to approve the official terms of the project and related definitions,

ensuring consistency among the various research units and keeping into account disciplinary and cultural differences.[3]

Research Questions

The InterPARES Project set out the following research questions for the Terminology Research Team:

- Is the term proposed specific to a field? If so, is its definition agreed upon in such field? If other definitions exist, how does the definition proposed relate to the others used? If the term is not specific to a field, is it a term in common usage or a neologism? How is its use justifiable in the context of the research?
- Is the term proposed used in other fields as well? If so, is its definition consistent across such fields? If not, what are the justifications for using one definition over another?
- Is the term used in several languages/traditions? If so, are the definitions consistent? If not, what are the justifications for using one definition over another?
- Is the term proposed consistent with the terms already used by the project? If so, does such inconsistency warrant a review of the already accepted terms in light of the new findings?
- Does the term express a concept that is already wholly or partial expressed by other already accepted terms or more appropriate terms?

Research Initiatives

To fulfil its mandate and address its research questions, the Terminology Team was tasked with developing a Register of terms and phrases to be selected from a corpus of InterPARES 1 and 2 documents and the earlier UBC Project Glossary.[4] Using the terms in the Register, the team was then tasked with developing the following terminological instruments:

- A Glossary that will provide logical or conceptual definitions of the words and phrases in the register as they are to be used for working purposes within the Project, to provide for consistency.
- A Dictionary that will provide discipline-specific logical or conceptual definitions of the terms or phrases in the glossary (as well as additional terms and phrases that are not in the glossary) as they apply to the various disciplines.
- A Thesaurus (i.e., a list of terms and their relationships).[5]

Collectively, these instruments were to comprise what became known as the InterPARES 2 Terminology Database.

[3] See Luciana Duranti (2001), "International Research on Permanent Authentic Records in Electronic Systems (InterPARES): Experiential, Interactive and Dynamic Records," SSHRC MCRI InterPARES 2 Project Proposal, 412-2001, 1.1-7. Available at http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf.
[4] See http://www.interpares.org/UBCProject/gloss.htm.
[5] This initiative began as a class project when a group of students at the University of California at Los Angeles, who had been working on a re-design of the US-InterPARES Web site, were encouraged by the then Terminology Team chair, Jonathan Furner, to build a thesaurus as part of the Web site redesign effort. The students began with the InterPARES Glossary and built a faceted, hierarchical structure on top, using the TCS-8 thesaurus construction tool available from WebChoir (http://www.webchoir.com/). Although never completed, a draft version of the thesaurus, which contains approximately 1,200 terms, is available online at http://polaris.gseis.ucla.edu/jfurner/IQ2/home.htm.

Research Methodology

The typical approach for populating the Terminology Database was for the terminology team to receive proposed terms from each of the Project's research units, together with proposed definitions. The terminology team's task was to then research the term according to the research questions listed above, by examining past InterPARES and UBC Project research documents, as well as the relevant dictionaries, glossaries and literature of the fields and countries represented in the Project. The team also consulted general dictionaries, where appropriate. On the basis of the result of such research, the team either: (1) accepted a term with its submitted definition, (2) returned it to the relevant research unit with proposed changes to the definition and/or a proposed alternate term or (3) rejected the term.

Terminology Database

Purpose

The original purpose of the InterPARES 2 Terminology Database was to support the Project's researchers in understanding concepts across disciplines that have come to address issues of preserving dynamic, experiential and interactive authentic digital records in electronic systems. Archival Science, Information Science, Computer Science, Geomatics, Music, Film, Dance, Law and several other disciplines all brought expertise and terminology to this Project. The Terminology Database facilitated communication and research among researchers in different disciplines and across cultural boundaries throughout the course of the Project by defining terms and relationships among terms used in the research. By extension, the Terminology Database now stands as a significant contribution to the understanding of records in dynamic, interactive and experiential systems in the artistic, scientific and governmental sectors.

Scope

To prepare the terms and the definitions for the Terminology Database, the team consulted resources relevant to each of the fields and disciplines represented in the Project. These included Arts (Film, Music, Dance), Sciences (Geomatics, Physics, Astronomy, Archaeology), Government (Law, Policy, E-Government) and Archival, Information and Computer Sciences. Terminology developed and used in InterPARES 1 and in the earlier UBC Project was carried forward to the InterPARES 2 Terminology Database. In this sense, the InterPARES 2 Project built upon its past knowledge, making it relevant to the contemporary environment.

Structure

There are two main components to the Terminology Database. They are (1) the Glossary and (2) the Dictionary. The Glossary is the authoritative list of terms and definitions that were core to the researchers' understanding of the evolving records creation, recordkeeping and records preservation environments. As the authoritative list, the Glossary provided the approved terms and definitions used throughout the InterPARES 2 research in both working documents and published documents. Because many, if not most, of the terms that were approved have multiple cross-disciplinary and/or cross-cultural definitions, it was important to account for plurality of meaning, which is precisely what the Dictionary is intended to do. The Dictionary contains all the terms from the Glossary, but, unlike the Glossary, the Dictionary provides multiple definitions for a single term gleaned from multiple disciplines. Each of these definitions are cited as coming from a particular discipline (e.g., Science, Archives or Arts), and from a particular source (e.g., Art and Architecture Thesaurus, Dictionary of Computing, or A Glossary of Archival and Records Terminology). Sources are of two kinds: dictionaries in the field or research documents from InterPARES 1 or 2. In short, the Dictionary is a tool used to facilitate interdisciplinary communication among members of the various communities that have a stake in the use of the terms used in the Project; for example, to support the writing of guidelines. The discipline-specific definitions of terms in the Dictionary has allowed the findings of the Project, many of which are cast in InterPARES-specific terminology, to be recast in the language of artistic, scientific and government disciplines. Thus, for example, by using this tool, Project researchers were able to see how Archival Science deploys terminology compared to researchers and practitioners in Computer Science, Library and Information Science, the Arts, etc., as demonstrated above in the introduction for the term "record."

A third terminological instrument, the Ontologies, was developed to identify explicit relationships among concepts. This instrument is particularly useful for communicating the nuances of Diplomatics in the dynamic, experiential and interactive environment.

All three of these instruments were drawn from a Register of terms gathered over the course of the Project. This Register served as a holding place for terms and phrases, and allowed researchers to discuss, comment on and modify submissions. The Register and the terminological instruments were housed in the online Terminology Database. The Database provides searching, display and file downloads, making it easy for users to navigate through the terminological instruments.

Statistics and structure of the terminological instruments

Dictionary

The Dictionary is the largest of the terminological instruments. It contains more than 900 terms, most of which contain multiple definitions. As shown in Figure 1, the entries in the Dictionary follow a standard format: term, part of speech, definition, disciplinary classification and citation. As well, it is important to note that, in cases where a Dictionary entry also appears in the Glossary, the Glossary definition is always the first of any multiple definitions listed for the entry in the Dictionary.

tool

n., Information technology, or other equipment or supply used to manage records at any time during their lifecycle. [Archives - *Manage Chain of Preservation,* Page: n.p.]

n., A program used primarily to create, manipulate, modify, or analyse other programs, such as a compiler or an editor or a cross-referencing program. Opposite: app, operating system. [Computer and Information Sciences - *Free Online Dictionary of Computing (FOLDOC)*]

n., Any existing physical object that is in some way fashioned or altered by humans and employed for a specific task or purpose. [Sciences - *Encyclopedic Dictionary of Archaeology*]

Figure 10. "Tool" entry in InterPARES 2 Project Dictionary

Glossary

The Glossary is smaller than the Dictionary because it includes the only meaning in which each term is used within InterPARES 2 documents. The Glossary terms with their definitions are the key to the communication of the findings of InterPARES 2. There are more than 600 terms in the Glossary (excluding terms and phrases that are specific to the Chain of Preservation and Business-driven Recordkeeping models). The format is similar to the Dictionary, but, as already stated, with only a single definition for each term. Exceptions arise when there is a single term for many concepts, such as in the example provided in Figure 2.

archives

n., [records] The whole of the documents made and received by a juridical or physical person or organization in the conduct of affairs, and preserved. [Archives - *School of Library, Archival and Information Studies (SLAIS) Glossary*]

n., [institution] An agency or institution responsible for the preservation and communication of records selected for permanent preservation. [Archives - *The InterPARES 1 Project Glossary*]

n., [place] A place where records selected for permanent preservation are kept. [Archives - *The InterPARES 1 Project Glossary* , Page: 356]

Figure 11. "Archives" entry in InterPARES 2 Project Glossary—illustrating a polysemous term

A key support mechanism for both the Glossary and Dictionary is the InterPARES 2 Bibliography. Where possible, the Project rooted its definitions in extant literature and cited the source for each definition that appears in the Terminology Database. This truly provides context for the Project's definitions because it stitches the use of the words and concepts as understood by the Project's researchers with the same words and concepts used in other texts by researchers and practitioners outside of the Project.

Ontologies

To illustrate explicit relationships among terms in the Glossary, the Terminology Cross-domain constructed several Ontologies. Within the context of the Project, the Ontologies were used to illustrate the relationship among Diplomatic terms. The terminology team constructed three Ontologies: (1) Archives and Records, (2) Status of Transmission and (3) Trustworthiness.[6] All of these Diplomatic concepts are core to the Project's understanding of records in dynamic, experiential and interactive systems and are made clearer through specific explication in the Ontologies. Figure 3 provides an example of an ontology by illustrating the component parts of Trustworthiness and their relationships.

Public access

At the close of the InterPARES 2 Project, the Terminology Database was "frozen" and is available to the general public via the InterPARES Web site.[7] The Dictionary and Glossary are searchable and downloadable in PDF format, and the Ontologies are available in PDF format.

[6] See Appendix 22.

[7] See http://www.interpares.org/ip2/ip2_terminology_db.cfm.

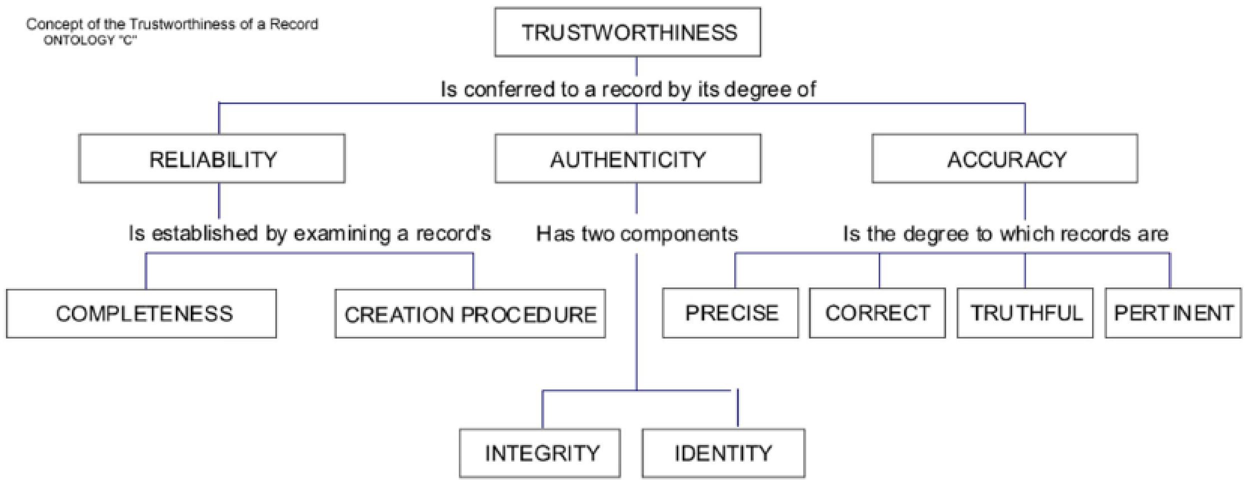


Figure 12. InterPARES 2 Project Ontology C - "Trustworthiness"

Conclusions

The Terminology Cross-domain Task Force acted as a service research team to the InterPARES 2 Project. It is hoped that it fulfilled the same role to Archival Science. In the process of constructing the three terminological instruments discussed above, the Project made great steps in advancing the general understanding of Archival concepts in the contemporary environment and expanded the existing records- and preservation-related vocabulary to include new ways of thinking about age-old problems.

Terminology work is not without its challenges. The challenges of terminological work—on any scale—are challenges of time and socio-political negotiation. As knowledge develops over time new terms surface, old terms are reinterpreted, forgotten, ossified in their original context, or brought forward into a new and vibrant scholarly discourse. Archival terminology is rich in expressiveness and history. As a service research team, the Terminology Cross-domain aimed at keeping terminological tradition, innovation and integration a positive and productive venture for the Project researchers and for Archival Science.

APPENDIX 1

InterPARES 2 Project Members

InterPARES 2 Project Members

Researchers:

Brad Abbott	Umgeni Water (South Africa)
Alicia Barnard	Ministry of Health, Mexico
Paul Berkman	University of California, Santa Barbara—Bren School of Environmental Science and Management/EvREsearch, Ltd.
Howard Besser	New York University—Cinema Studies Department
Karuna Bhoday	Monash University—School of Information Management and Systems
Justine Bizzocchi	Banff Centre—New Media Institute
Richard Blake	The National Archives of the United Kingdom
Jean-François Blanchette	University of California, Los Angeles—Department of Information Studies
Filip Boudrez	City Archives of Antwerp—The David Project
Marta Braun	Ryerson University—School of Image Arts
Paolo Buonora	State Archives of Rome
Greg Burley	Carnegie Observatories (MOST Satellite Mission)
Ann Butler	New York University—Fales Library
Margaret Campbell	Archives of Nova Scotia
Martine Cardin	Université Laval—Institut sur le patrimoine culturel
Su-Shing Chen	University of Florida—Computer, Information Science and Engineering
Wei Chen	Beijing Municipal Archives
Hao Chenhui	State Archives Administration of China
Michèle Cloonan	Simmons College—Graduate School of Library and Information Science
Barbara Craig	University of Toronto—Faculty of Information Studies
Henry Daniel	Simon Fraser University—School for the Contemporary Arts
Ben Howell Davis	Davis International Associates
Vincenzo De Meo	Associazione Nazionale Archivistica Italiana
Hannelore Dekeyser	Katholieke Universiteit Leuven—Faculty of Law—Interdisciplinary Centre for Law and Information Technology
Glenn Dingwall	City of Vancouver Archives
Luciana Duranti	The University of British Columbia—School of Library, Archival and Information Studies
Terry Eastwood	The University of British Columbia—School of Library, Archival and Information Studies
Fynnette Eaton	United States National Archives and Records Administration
Philip Eppard	University at Albany, State University of New York—Department of Information Studies
Joanne Evans	Monash University—School of Information Management and Systems
Sharon Farb	University of California, Los Angeles—Libraries

Sidney Fels	The University of British Columbia—Department of Electrical and Computer Engineering
Gigliola Fioravanti	Associazione Nazionale Archivistica Italiana
Arianna Franceschini	Associazione Nazionale Archivistica Italiana
Jonathan Furner	University of California, Los Angeles—School of Information Management and Systems
Chris Gallagher	The University of British Columbia—Department of Theatre and Film
Victoria Gebert	Library and Archives Canada
Mahnaz Ghaznavi	J. Paul Getty Trust—Institutional Records and Archives
Anne Gilliland	University of California, Los Angeles—Department of Information Studies
Kevin Glick	Yale University Archives
Elaine Goh	National Archives of Singapore
Arien Crespo Gonzáles	Casa de las Américas (Cuba)
Maria Guercio	Universita' degli Studi di Urbino—Istbal
Yvette Hackett	Library and Archives Canada
Keith Hamel	The University of British Columbia—School of Music
Babak Hamidzadeh	Library of Congress
Ken Hannigan	National Archives of Ireland
Chenhui Hao	State Archives Administration of China—Institute of Scientific & Technical Research on Archives
P.C. Hariharan	Systems Engineering & Security, Inc.
Kenneth Hawkins	United States National Archives and Records Administrations
Graeme Hirst	University of Toronto—Department of Computer Science
Hans Hofman	National Archives of the Netherlands
Wang Hongmin	State Archives Administration of China
Holger Hoos	The University of British Columbia—Computer Science
Sally Hubbard	Getty Research Institute—Getty Standards and Digital Resource Management Program
Margaret Hutchison	British Columbia Archives
Livia Iacovino	Monash University—Information Technology
Mary Ide	WGBH Educational Foundation
Glen Isaac	British Columbia Archives
Yang Jibo	State Archives Administration of China
Michael Joyce	Vassar College—English Department
Agnes Jonker	Netherlands Archives School
Andrea Kalas	DreamWorks SKG—Archives Division
Wayne Kelly	Earthstar Communications
Rick Kopak	The University of British Columbia—School of Library, Archival and Information Studies
Claudia Lacombe	National Archives of Brazil
Wang Lan	State Archives Administration of China
Ian Lancashire	University of Toronto—Department of English
Brent Lee	University of Windsor—School of Music
Jiang Li	State Archives Administration of China

Wang Liangcheng	State Archives Administration of China
Yu Lijuan	Renmin University of China—Archives
Michael Longton	University of Victoria—School of Music
Warren Lowe	Nucleus Labs
Randal Lukow	DreamWorks SKG—Archives Division/Turner Broadcasting Systems Archives
Richard Lysakowski	Collaborative Electronic Notebook Systems Association
Bonnie Mak	University of Illinois—Graduate School of Library and Information Science
Alessio Malizia	University of Rome—Computer Science Department
Timothy Maloney	University of Minnesota—Music Library
Richard Marciano	San Diego Supercomputer Center
Enrica Massella	Autorità per l'Informatica nella Pubblica Ammistrazione
Terrence Maxwell	University at Albany, State University of New York—Department of Information Studies
Vicky McCargar	McCargar Consulting
John McDonough	National Archives of Ireland
Sue McKemmish	Monash University—School of Information Management and Systems
Evelyn McLellan	Insurance Corporation of British Columbia Archives
Du Mei	State Archives Administration of China
Aída Mendoza	Consultant (Péru)
Mattia Miani	Legacoop of Bologna (Italy)
Reagan Moore	San Diego Supercomputer Center
Michael Murphy	Ryerson University—Rogers Communications Centre
Vivek Navale	United States National Archives and Records Administration
Isabella Orefice	Associazione Nazionale Archivistica Italiana
Catherine Owen	University of Glasgow—Performing Arts Data Service, Department of Theatre, Film and Television Studies
Giovanni Paolini	Università La Sapienza, Roma—Scuola Speciale per Archivisti e Bibliotecari
Eun Park	McGill University—Graduate School of Library and Information Studies
Richard Pearce-Moses	Arizona State Library—Archives and Public Records
Barbara Pernici	Politecnico di Milano (Italy)
Cristine Petillat	National Archives of France
Jean-Stéphen Piché	Library and Archives Canada
Cecilia Pirola	Associazione Nazionale Archivistica Italiana
Fosca Pizzcaroni	Associazione Nazionale Archivistica Italiana
Laura Primangeli	Professional photographer, Italy
Xiaowei Qiu	State Archives Administration of China
Tom Quinlan	National Archives of Ireland
Jesse Read	The University of British Columbia—School of Music
Geoffrey Rockwell	McMaster University
Andrew Rodger	Library and Archives Canada—Art and Photography Section
John Roeder	The University of British Columbia—School of Music

Rosely Rondinelli	Museu do Índio—FUNAI (Brazil)
Shelby Sannett	United States National Archives and Records Administration
Andrew Schloss	University of Victoria—School of Music
Guo Siping	State Archives Administration of China
Leon Stout	Penn State University—University Archives
Jim Suderman	Provincial Archives of Ontario and, later, City of Toronto Archives
Anna Szlejcher	National University of Cordoba (Argentina)—School of Archivología
Hiroya Takeuchi	Shizuoka University of Art and Culture—Faculty of Cultural Policy and Management
Fraser Taylor	Carleton University—Department of Geography & Environmental Studies
Joe Tennis	The University of British Columbia—School of Library, Archival and Information Studies/University of Washington—Information School
Enrica Massella Ducci Teri	Autorità per l'Informatica nella Pubblica Ammistrazione (Italy)
Kate Theimer	United States National Archives and Records Administration
Kenneth Thibodeau	United States National Archives and Records Administration
Lynne Tibbitt	British Columbia Archives
Malcolm Todd	The National Archives of the United Kingdom
Ciaran Trace	University of California, Los Angeles—Department of Information Studies
James Turner	Université de Montréal—École de bibliothéconomie et des sciences de l'information
William Underwood	Georgia Tech Research Institute
Bruce Walton	Library and Archives Canada
Ronald Weiss	Arkival Technology Corporation
Adrienne Woods	United States National Archives and Records Administration
George Wootton	Earthstar Communications Group
Peter Van Garderen	Artefactual Systems, Inc.
Stefano Vitali	State Archives of Florence (Italy)
Ma Xueqiang	Archives of the Shandong Province of China
Shuxia Zhang	State Archives Administration of China
Zhao Zhongxin	State Archives Administration of China

Research Assistants:

Melissa Adams	The University of British Columbia—MAS program
Scott Amort	The University of British Columbia—PhD School of Music
Tom Anderson	The University of British Columbia—MAS program
Jennifer Baetz	The University of British Columbia—MAS program
Bart Ballaux	The University of British Columbia—MAS program
Gary Barclay	The University of British Columbia—MAS program
Patsy Baudoin	Simmons College—Graduate School of Library and Information Science

Barbara Bean	University at Albany, State University of New York—Department of Information Studies
Tiran Behrouz	The University of British Columbia—BSc Computer Science
Lindsay Bergen	The University of British Columbia—MAS program
David Boudinot	The University of British Columbia—MAS/MLIS program
Deidre Brocklehurst	The University of British Columbia—MAS program
Jessica Bushey	The University of British Columbia—MAS program
Lauren Cardinal	University at Albany, State University of New York—Department of Information Studies
Wyndy Carr	The University of British Columbia—MLIS program
Carolyn Caseñas	The University of British Columbia—MAS/MLIS program
Natalie Catto	The University of British Columbia—MAS/MLIS Program
May Chan	The University of British Columbia—MAS/MLIS Program
Chia-Ning Chiang	The University of British Columbia—PhD program
Yau Min Chong	The University of British Columbia—MAS program
Luca Ciccarelli	Politecnico de Milano
Erin Coulter	The University of British Columbia—MAS program
Luca Ciccarelli	Politecnico di Milano
Naomi Cull	University of Toronto—PhD Department of Linguistics
Seth Dalby	The University of British Columbia—MAS program
Heather Daly	The University of British Columbia—MAS/MLIS program
Kimberley Davison	The University of British Columbia—MAS/MLIS program
Heather Dean	The University of British Columbia—MAS/MLIS program
Terra Dickson	The University of British Columbia—MAS program
Glenn Dingwall	The University of British Columbia—MAS program
Jennifer Douglas	The University of British Columbia—MAS and PhD programs
Alan Doyle	The University of British Columbia—MAS program
Joanne Evans	Monash University—Information Technology
Cobi Falconer	The University of British Columbia—MAS program
Adam Farrell	The University of British Columbia—MAS/MLIS program
Dan Farrell	The University of British Columbia—MAS program
Ann Forman	The University of British Columbia—MAS program
Fiorella Foscarini	The University of British Columbia—PhD program
Shanna Fraser	The University of British Columbia—MAS program
Stephen Gage	The University of British Columbia—MAS program
Peter Gagné	Université Laval—Institut sur le patrimoine culturel
Michael Garabedian	University of California, Los Angeles—School of Information Management and Systems
David Gibbs	University of California, Los Angeles—School of Information Management and Systems
Jessica Glidewell	The University of British Columbia—MAS program
Monica Greenan	The University of British Columbia—MAS program
Susan Gutmann	Universtiy of British Columbia—MLIS program
Nadine Hafner	The University of British Columbia—MAS program
Erin Hanlon	The University of British Columbia—MAS/MLIS program

Joshua Hauck-Whealton	University at Albany, State University of New York—Department of Information Studies
Keum Hee	The University of British Columbia—MAS program
Peggy Heger	The University of British Columbia—MAS/MLIS program
Sarah Henshaw	The University of British Columbia—MAS program
Ted Hoppenstedt	University at Albany, State University of New York—Department of Information Studies
Janine Johnston	The University of British Columbia—MAS program
John Juricek	University of California, Los Angeles—Department of Information Studies
Sharif Khandaker	The University of British Columbia—MAS program
Eleanor Kleiber	The University of British Columbia—MAS/MLIS program
Greg Kozak	The University of British Columbia—MAS program
Tracey Krause	The University of British Columbia—MAS/MLIS program
Andrea Lam	The University of British Columbia—MAS/MLIS program
Karen Langley	The University of British Columbia—MAS program
Alison Langmead	University of California, Los Angeles—Department of Information Studies
Tracey Lauriault	Carleton University—PhD Department of Geography and Environmental Studies
Monique Leahey-Sugimoto	University of California, Los Angeles—Department of Information Studies
Lori Lindberg	University of California, Los Angeles—PhD Department of Information Studies
Megan Lindenbach	The University of British Columbia—MAS program
David Litke	The University of British Columbia—PhD program
Yvonne Loiselle	The University of British Columbia—PhD program
Anne MacLean	The University of British Columbia—MAS program
Scott McAlpine	The University of British Columbia—MAS program
Susan McGoey	The University of British Columbia—MAS program
Rachel McMullin	University at Albany, State University of New York—Department of Information Studies
Brenda McPhail	University of Toronto—Faculty of Information Studies
Luke Meagher	The University of British Columbia—MAS/MLIS program
Jennifer Meehan	The University of British Columbia—MAS program
Catherine Miller	The University of British Columbia—MAS program
Christina Miller	The University of British Columbia—MAS program
Katherine Miller	The University of British Columbia—MAS/MLIS program
Rachel Mills	The University of British Columbia—MAS program
Shaunna Moore	The University of British Columbia—MAS program
Jane Morrison	The University of British Columbia—MAS program
Elisheba Muturi	The University of British Columbia—MAS/MLIS program
Greg Myers	The University of British Columbia—MLIS program
Linda Nobrega	The University of British Columbia—MAS program
Erin O’Meara	The University of British Columbia—MAS program
Emily O’Neill	The University of British Columbia—MAS/MLIS program

Jennifer Osorio	University of California, Los Angele—Department of Information Studies
Heejin Park	The University of British Columbia—PhD program
Cara Payne	The University of British Columbia—MAS program
Philippe Perron	Université Laval—Institut sur le patrimoine culturel
Carolyn Petrie	The University of British Columbia—MAS/MLIS program
Randy Preston	The University of British Columbia—MAS program
Jill Pryor	The University of British Columbia—MAS program
Hemalatha Ramasamy	The University of British Columbia—MAS program
Marilyn Ramen	The University of British Columbia—MAS/MLIS program
Scott Redgrove	The University of British Columbia—MAS/MLIS program
Lisa Ricciuti	The University of British Columbia—MAS/MLIS program
Claudette Rocan	The University of British Columbia—MAS/MLIS program
Corinne Rogers	The University of British Columbia—MAS program
Molly Rooney	The University of British Columbia—MLIS program
Nadav Rouche	University of California, Los Angeles—Department of Information Studies
Peter Runge	University at Albany, State University of New York—Department of Information Studies
Rebecca Russell	The University of British Columbia—MAS program
Vince Schillaci-Ventura	The University of British Columbia—MLIS program
Anthea Seles	The University of British Columbia—MAS program
Geneviève Shepherd	The University of British Columbia—MAS and PhD programs
Arilea Sill	The University of British Columbia—MAS program
Julie Simard	Université Laval—Institut sur le patrimoine culturel
Michael Skipper	United States National Archives and Records Administration
Frédéric Smith	Université Laval—Institut sur le patrimoine culturel
Wendy Sokolon	The University of British Columbia—MAS/MLIS program
Emily Staresina	University of California, Los Angeles—Department of Information Studies
Stuart Sugarbread	University of California, Los Angeles—Department of Information Studies
Mary Beth Sullivan	University at Albany, State University of New York—Department of Information Studies
Shannon Supple	University of California, Los Angeles—Center for Information as Evidence
Melissa Taitano	University of California, Los Angeles—Department of Information Studies
Jill Teasley (Prior)	The University of British Columbia—MAS/MLIS program
Adele Torrance	The University of British Columbia—MAS program
Brian Trembath	The University of British Columbia—MAS/MLIS program
Melanie Wallace	The University of British Columbia—MAS program
Holly Wang	University of California, Los Angeles—PhD Department of Information Studies
Carol Ward	University at Albany, State University of New York—Department of Information Studies

Paul Washington	The University of British Columbia—MAS program
Reginald White	University at Albany, State University of New York—Department of Information Studies
Sherry Xie	The University of British Columbia—MAS and PhD program.
Catherine Yasui	The University of British Columbia—MAS program
Eunha Youn	University of California, Los Angeles—Department of Information Studies
Keum Hee Yu	The University of British Columbia—MAS program
Jessica Zacher	University at Albany, State University of New York—Department of Information Studies
Yuchai Zhou	Carleton University—Department of Geography and Environmental Studies

Advisory Board Members:

Chris Hurley	Commonwealth Bank (Australia)
Shadrack Katuu	Nelson Mandela Foundation (South Africa)
Isabelle de Lamberterie	Centre National de la Recherche Scientifique (CNRS) (France)
S. Timothy Maloney	University of Minnesota (USA)
John McDonald	Consultant (Canada)
Donald Sawyer	NASA/Science Office of Standards and Technology (NOST) (USA)

Post-doctoral Fellows:

Jean-François Blanchette	2002-2004
Brent Lee	2002
Bonnie Mak	2004-2006

Project Management:

Yau Min Chong	Project Coordinator, 2003-2005
Tahra Fung	Project Coordinator, 2002-2002
Greg Kozak	Project Coordinator, 2005-2006
Jean-Pascal Morghese	Technical Coordinator, 2002-Present
Sandy Orr	Project Assistant, 2002-Present
Randy Preston	Project Coordinator, 2006-Present
Mark Wolfe	Project Manager (USA), 2002-2006

APPENDIX 2

The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES[†]

by

Luciana Duranti, The University of British Columbia
Kenneth Thibodeau, National Archives and Records Administration

[†] First published in *Archival Science* (2006) 6(1): 13-68. Reprinted with permission of the publisher.

Archival Science (2006) 6:13–68
DOI 10.1007/s10502-006-9021-7

The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES★

LUCIANA DURANTI[1] and KENNETH THIBODEAU[2]
[1]*The University of British Columbia, Vancouver, BC, Canada (E-mail: luciana@interchange.ubc.ca);* [2]*National Archives and Records Administration, Washington, DC, USA (E-mail: ken.thibodeau@nara.gov)*

Abstract. This article presents the concept of electronic record as articulated by the first phase of the InterPARES (International research on Permanent Authentic Records in Electronic Systems) Project (1999–2001) and discusses it in light of the findings of the second phase of the Project (2002–2006). While InterPARES 1 focused on records produced and/or maintained in databases and document management systems, InterPARES 2 examined records produced and/or maintained in interactive, experiential and dynamic environments. The authors describe the characteristics of these environments and of the entities found in them in the course of case studies conducted on systems used for carrying out artistic, scientific and e-government activities, and propose the new concept of record that InterPARES 2 is eleborating, which expands on that formulated by InterPARES 1.

Keywords: archival theory, concept of record, digital records, diplomatics, dynamic systems, e-government records, electronic art, electronic science, electronic systems, experiential systems, interactive systems

The InterPARES (International research on Permanent Authentic Records in Electronic Systems) Project aims at developing the theoretical and methodological knowledge essential to the long-term preservation of authentic records created and/or maintained in digital form. This knowledge should provide the basis from which to formulate model policies, strategies and standards capable of ensuring the longevity of such material and the ability of its users to trust its authenticity. InterPARES has developed in two phases. InterPARES 1, which ran from 1999 to 2001, dealt with records mandated for accountability and administrative needs that are created in databases and document management systems. InterPARES 2, which began in 2002 and will be completed by the end of 2006, has focused on the portion of society's

★ The authors would like to thank InterPARES 2 research assistant Randy Preston for his careful editing and constructive criticisms. Some of his suggestions have been incorporated in the text of this article.

recorded memory that is digitally produced in interactive, dynamic and experiential systems in the course, and as a byproduct, of artistic, scientific and electronic government activities.[1] The distinctive and novel characteristics of the environments being examined in InterPARES 2 force a re-examination of the findings of InterPARES 1, including its interpretation of traditional archival concepts, especially that of record. This article presents the work of two of the InterPARES 2 researchers. While the authors regard the findings of this article as conclusive, the article should be seen as a contribution to the ongoing InterPARES work, and not as a final product of the project.

This article examines the characteristics of documents in interactive, dynamic and experiential systems, as observed in InterPARES 2 case studies and other empirical instances, in order to determine whether they are or can be records and whether records in such systems have unique characteristics which might necessitate a revision of the traditional concept of a record. The analysis of record characteristics is based on the definition of 'record' from archival science and more specifically on its articulation in InterPARES 1.[2]

This analysis is divided into five sections. Section 1 summarizes relevant findings of InterPARES 1. Section 2 describes interactive, experiential and dynamic environments, and considers their implications for records made or received and/or kept in such environments. Section 3 focuses more specifically on the "documents" that are or might be created in such environments. Section 4 considers "records" that are or might be made or received and/or kept in these environments. Section 5 examines the keeping of such records. Finally, a concluding section draws out the major findings of this analysis.

---

[1] See the InterPARES website at http://www.interpares.org.

[2] There have been numerous projects that have examined the characteristics of digital documents. Most notable is the Open Archival Information System (OAIS) Reference Model, available at http://www.ccsds.org/docu/dscgi/ds.py/Get/File-143/650x0b1.pdf. The information model articulated in the OAIS standard has been the foundation of analysis of the characteristics of digital documents in several other projects, such as the CEDARS, PREMIS and Persistent Archives projects, respectively accessible at http://www.leeds.ac.uk/cedars/, http://www.oclc.org/research/projects/pmwg/, and http://www.sdsc.edu/NARA/. However, these projects have developed their characterizations of digital documents with a view towards addressing the practical challenges of preserving them. In contrast, this article seeks to identify the characteristics of digital documents that are records. The goal is to describe these records in themselves. Such description must be independent of and transcend any and all approaches to preserving them.

Findings of InterPARES 1

At the beginning of InterPARES 1, the research team adopted a concept and a definition for the terms record, document, information and data, and used them to identify the digital objects in the systems examined in its case studies. The team adopted the traditional archival definition of a record as any document created (i.e., made or received *and* set aside – i.e. kept, saved – for action or reference) by a physical or juridical person in the course of a practical activity as an instrument and by-product of such activity. It defined 'document' as recorded information, 'information' as a message intended for communication across space or time, and 'data' as the smallest meaningful piece of information. Finally, an 'electronic record' was defined as a record that is set aside and used in electronic form irrespective of the original form in which it may have been made or received. The InterPARES 1 conception was fully consistent with the archival principle that whatever the creator treats as a record in the course of any given action is indeed a record in the context of this action. It also made it clear that what distinguishes a record from a document that is not a record is the nature of its relationship with the activity of the creator rather than its formal or content characteristics.

InterPARES 1 researchers explored the assumptions and implications of the definition of 'record' by determining what the necessary characteristics of an electronic record are on the basis of both archival theory and diplomatic theory. These two theories were regarded as complementary because, while diplomatic theory examines records as items, enabling identification of the characteristics embedded in the records themselves, archival theory, by treating records as parts of aggregations, examines their relationships to other records, to the persons involved in their creation, and to the activities in the course of which they are created and used. The research team identified the following necessary characteristics: (1) a fixed form, meaning that the entity's content must be stored so that it remains complete[3] and unaltered, and its message can be rendered with the same documentary form it had when

[3] Completeness here is not mentioned as a characteristic of the record, because an incomplete record is still a record, albeit a bad one, but as a characteristic of a fixed form, according to which a form that is fixed is one that does not lose any of its original elements in the process of being stored and retrieved.

first set aside; (2) an unchangeable content;[4] (3) explicit linkages to other records within or outside of the digital system, through a classification code or other unique identifier;[5] (4) an identifiable administrative context; (5) an author, an addressee, and a writer; and (6) an action, in which the record participates or which the record supports either procedurally or as part of the decision making process.[6]

Having specified the necessary characteristics of an electronic record, the research team accepted as a working hypothesis the fundamental assumption of diplomatics that, regardless of differences in nature, provenance or date, from a formal point of view all records are similar enough to make it possible to conceive of one typical, ideal documentary form containing all possible elements of a record. From this hyposthesis, the team derived the corollaries that, while they may manifest themselves in different ways, the same formal elements that are present in traditional records exist in electronic records either explicitly or implicitly, and that all electronic records share the same formal elements. Thus, the team created a template, that is, a decomposition of the ideal electronic record, first, into its constituent parts,

[4] The stability of the record, as determined by its fixed form and its unchangeable content, is only implied in the part of the archival definition that reads that a record is a document (i.e., rather than just data or information), but it is explicitly stated in the diplomatic definition and concept of record. (See Luciana Duranti, *Diplomatics. New Uses for an Old Science* (Lanham, Maryland, and London: The Scarecrow Press, Inc., The Society of American Archivists and the Association of Canadian Archivists, 1998), pp. 41–58.

[5] This characteristic corresponds to the archival bond, which is implied in the archival definition when records creation is linked to an activity, but it is made explicit by archival theorists of all times and cultures. See Luciana Duranti, "The Archival Bond," *Archives and Museum Informatics* 11, nos.3–4 (1997): 213–218.

[6] While characteristics 4 and 6 can be deduced from the archival definition, characteristic 5 derives from the diplomatic concept of record: it was considered important in order to distinguish records from digital objects resulting from simply querying a database. The author is the person issuing the record, the writer is the person determining the articulation of the discourse in the record, and the addressee is the person for whom the record is intended. As a record must participate in an action and any action must fall on somebody, the addressee is necessary to the existence of the record. See the Appendix 2 of the book *The Long-term Preservation of Electronic Records: the InterPARES Project* on the InterPARES website at http://www.interpares.org/book/index.cfm. This book is also distributed in print by the Society of American Archivists: Luciana Duranti ed., *The Long-term Preservation of Electronic Records: the InterPARES Project* (San Miniato, Italy: Archilab, 2005).

and then, within the part "form," into its elements.[7] In the template, the parts and elements are defined and their purpose is explained. The research team used the template as an instrument for the systematic analysis of the electronic objects contained in several different systems, for the purpose of establishing which ones are records.

The template is composed of four sections corresponding to the four necessary constituent parts of every record: documentary form, annotations, context, and medium.[8] The documentary form[9] includes, among the intrinsic elements,[10] the names of the persons concurring to the creation of the record, the chronological date, the place of origin of the record, the indication and description of the action or matter, the attestation, and a statement of validation; and, among the extrinsic elements,[11] overall presentation features (e.g. text, image, sound, graphic), specific presentation features (e.g. layouts, hyperlinks, colors, sample rate of sound files, resolution of image files, scales of maps), electronic signatures and seals (e.g. digital signatures), digital time stamps, and special signs (e.g. digital watermarks, organization crests, personal logos).[12]

---

[7] The reason why the constituent parts of the record ended up in the template that is supposed to represent the ideal form of a record is that all identified constituent parts used to be regarded as necessary extrinsic elements of form by traditional diplomatists. It was important to show their presence, definition and purpose, and the fact that they are now independent of form.

[8] In a previous research endeavour commonly known as the UBC-DoD project, the parts constituting the records were identified as: medium, form, action, persons, archival bond, content and context. See Luciana Duranti and Heather MacNeil, "The Preservation of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project," *Archivaria* 42 (Spring 1997): 46–67; and Luciana Duranti, Terry Eastwood and Heather MacNeil, *Preservation of the Integrity of Electronic Records* (Dordrecht: Kluwer Academic Publishers, 2002): Chapter 1. In the context of InterPARES, it was decided that action, persons, archival bond and content, contrary to the other constituent parts, continue to manifest themselves in formal elements and are inextricable from them, so they do not have to be identified separately from the form. As it regards the annotations, which in the UBC-DoD project were included among the elements of form, they were added to the constituent parts because they are often linked to the record rather than embedded in it, and need therefore to be looked at separately from the record form.

[9] Defined as "The rules of representation according to which the content of a record, its administrative and documentary context, and its authority are communicated."

[10] Defined as "The elements of a record that convey the action in which the record participates and its immediate context."

[11] Defined as "The elements of a record that constitute its external appearance."

[12] See "Template for Analysis,"Appendix 1 in *The Long-term Preservation of Electronic Records: the InterPARES Project*, cit.

The annotations[13] fall into three fundamental groups: (1) additions made to the record after its creation as part of its transmission (e.g. priority of transmission, date of compilation and date of transmission in an e-mail record, the indication of attachments), (2) additions made to the record in the course of handling the business matter in which the record participates (e.g. date and time of receipt, action taken, name of handling office), and (3) additions made to the record in the course of managing it as a record (e.g. filing date, class code, registration number). The categorization of the contexts of the record[14] and the list of what would reveal them correspond to an hierarchy of frameworks that goes from the general to the specific: (1) juridical-administrative context (manifested in, for example, laws and regulations), (2) provenancial context (manifested in, for example, organizational charts, annual reports, tables of users in a database), (3) procedural context (manifested in, for example, workflow rules, codes of administrative procedure), (4) documentary context (manifested in, for example, classification schemes, records inventories, indexes, registers), and (5) technological context (manifested in, for example, hardware, software, system models, system administration).[15]

The medium[16] was difficult to place within the template, because, although it is still necessary for an electronic record to exist, it is not inextricably linked to the message, does not store the record as such, but stores one or more bit-streams which can be used to reproduce the record, and its choice by the record-maker or keeper can be either arbitrary or based on reasons related to preservation rather than to the function of the record. In addition, the medium is not a relevant factor in assessing a record's authenticity – one of the primary purposes of InterPARES – at least from the perspectives of the creator and of the record preserver.[17] This was confirmed by the case studies undertaken by the research team, by the end of which the team was convinced that, with electronic records, the medium should not be considered a constituent part of the record but a part of the record technological context.

[13] Defined as "Additions made to a record after it has been created."

[14] Defined as "The framework of action in which the record participates."

[15] For details related to annotations and contexts, see the Template for Analysis referenced above.

[16] Defined as "The physical carrier of the message."

[17] An additional reason for the InterPARES team to dissect the concept of record was to identify what parts or elements contribute to the authenticity of the record and to the ability to verify it.

While the physical medium is not a significant consideration for electronic records, the way an electronic record is organized and stored in one or more bit streams has a role analogous to that of physical medium for analog records. The research team identified those bit strings necessary to reproduce an electronic record and requiring distinct preservation measures as its *digital components*. With an analog record, the choice of physical medium is dependent on the overall presentation features of the record: a textual record can be recorded on paper, but an audio record cannot. With an electronic record, its organization into one or more digital components is dependent on the data type of the record. 'Data type' is a set of binary values used to encode data. Textual data can be encoded as ASCII, but audio data cannot. An analog record may be preserved authentic even when it is copied from one physical medium to another, provided the replacement medium is appropriate; for example, textual records originally recorded on paper can be preserved as authentic copies on microfilm. Similarly, electronic records can be authentically preserved even when they are transformed from one set of digital components to another, provided the replacement set preserves all the essential attributes of the record.

The concept of digital component was elaborated as a consequence of the recognition that it is literally not possible to preserve an electronic record like a record on paper. An electronic record is an object that is output from a computer system, typically on a screen, when needed by a human, or in interactions between systems, but cannot be stored in the form in which it is seen or used,[18] except by being converted to an analog form outside the system, but in that case it would no longer be an electronic record. Instead, it is stored as one or more strings of bits that require processing by a computer to be seen or used again as a unit. Thus, the research team determined that, empirically, preserving an electronic record consists of preserving the ability to reproduce it. A system that preserves electronic records must be able to identify and locate all the digital components of each record and apply the appropriate software to each component to reproduce the record.

[18] There may be only minor differences between the form in which a record used in interactions between systems, rather than presented to a human, is stored on a digital medium and the form in which it is used in automated transactions; nevertheless, it remains true that the way the digital data which constitute the record are inscribed on a physical medium, any physical medium, and the form in which they are transmitted between systems or the form in which they are stored in a computer's memory during transactions are never identical. In contrast, a traditional, analog record is inscribed on paper and transmitted and read in exactly that form.

Digital components may contain all or part of a record, and/or the related metadata. For example, an e-mail containing a textual message, a picture and a digital signature has at least four digital components: the header data, which enable systems to properly route and manage the message, the text of the message, the picture, and the digital signature. In contrast, a report with four textual attachments might be constituted of only one '.pdf' file (i.e., one digital component), but it might also consist of a word processing file that contains the body of the report, and four other files, possibly in different formats, for each of the attachments (i.e., five digital components). An important aspect of digital preservation is that it is possible to preserve the ability to reproduce an electronic record even when its digital components are altered. A report consisting of five word processing files could be combined into one file, and then converted from a word processing format into '.pdf.' So long as a '.pdf' reader faithfully renders the same document as would have been displayed by the original word processing software to the original five files, it would not matter that the encoding of the record in digital components had changed from five files to one and from a word processing format to '.pdf.'

Finally, the InterPARES 1 team felt the need to point out that the relation between a electronic record and a computer file can be one-to-one, one-to-many, many-to-one, or many-to-many, thus we should never use the terms record and file interchangeably; that the same presentation[19] of a record can be created by a variety of digital presentations and, vice versa, from one digital presentation a variety of record presentations can derive, thus fixed form does not imply that the bit streams must remain intact over time; and that it is possible to change the way a record is contained in a computer file without changing the record, thus the name of a record's documentary form does not necessarily indicate what digital object we are dealing with.

The analyses of the case studies conducted using the template indicated that only about one half of the examined systems contained records (12 out of 22), primarily because the objects identified within the other half did not appear to possess either a fixed documentary form or a stable content. When systems did contain records, these could rarely be compared with the model represented by the template, because, although they were able to achieve their purposes, they were not good records. For example, in most systems, there was no explicit

[19] 'Presentation' in this analysis means the act of materializing the overall and specific presentation features of an electronic record or the result of this action.

manifestation of the relationship among the records participating in the same affair or matter, and, although it was easy to identify the business processes supported by the system, it was not always possible to determine how the records participated in or supported specific actions. In addition, it was often difficult to determine the significance of the presence or absence of given elements of documentary form or of annotations.

More importantly, the case studies showed that, with electronic records, a key concept to consider is that of *records attributes*, which are the defining characteristics of each given record or of a record element in it. A *record element* is a constituent part of the record's documentary form and, as seen earlier, may be either extrinsic, like a seal, or intrinsic, like the salutation.[20] An attribute may manifest itself in one or more record elements. For example, the name of the author of a record is an attribute, which may be expressed as a letterhead or a signature, both of which are intrinsic elements of documentary form, that is, record elements. In addition to attributes that manifest themselves in the form of the record, that is, on the face of the record, as record elements, every record has attributes that are implicit elsewhere, such as the name of the creator or of the medium, but in electronic records they are explicit, albeit expressed outside the documentary form. They are mostly transparent to the user, because they manifest themselves as metadata included in either a record profile,[21] another digital object linked to the record, or documentation about the system or application in which the record is created. Attributes made explicit outside the record as metadata demonstrating its identity are *important* to uniquely identify any electronic record, but they are *essential* to the identification of digital objects that do not have – or at least for as long as they lack – a stable content or a fixed form.

The concept of electronic record presented above, with all its characteristics, parts, formal elements, attributes and digital components, has worked quite well with databases and document management systems. However, it may appear problematic when applied to the objects examined by InterPARES 2, because fluidity is part of their nature and contributes to the accomplishment of their purpose as instruments of, or support for action.

---

[20] A defining characteristic, or attribute, of the record element "seal" may be its legend.

[21] A record profile is an annotation inextricably linked to the record that includes several fields, which are either automatically or manually filled in with the record's metadata.

Interactive, Experiential and Dynamic Environments

InterPARES 2 has conducted case studies in the artistic,[22] scientific, and electronic government fields, focusing on cases which make use of digital technologies in innovative ways. This focus enables the project to examine whether there are differences in the nature of the records produced in environments that only exist in the digital domain. For the purpose of beginning its investigation of new technologies, InterPARES 2 initially adopted the Institute of Electrical and Electronics Engineers (IEEE) definition of an interactive system as "one in which each user entry causes a response from or an action by the system."[23] Interactive systems include a wide range of possibilities, starting with simple cases where the user's entries are constrained to a limited number of choices and the system's responses follow fixed paths from these choices. More complex situations occur where the number and variety of possible interactions are so great that the results are practically unpredictable. In even more complex cases, user input may become or generate new data that are included in subsequent outputs. The systems examined by InterPARES 2 include more complex and less predictable sequences of interactions with users and applications; for example, in the Electronic Café International[24] system, a multimedia international network for showcasing creative, multicultural, multidisciplinary, collaborative telecommunications, actions or responses are triggered by inputs from other systems.

There is a large class of applications where a system carries out individual transactions acting not simply as a machine, but as an agent for the system owner. This class includes systems used for electronic funds transfers between financial institutions and, more commonly, automated teller machines (ATM). Such systems produce, for the bank or other financial institution, an electronic record of a withdrawal, deposit, or transfer of funds, but do so without any physical,

[22] Traditionally, works of art and recordings of performance art would be regarded as end products of artistic activities, rather than records. However, the traditional definition of 'record' in archival science poses no restriction on the type of information object which may be a record. Provided it satisfies the requirements for records summarized in section 1 above, a work of art or recording of an artistic performance may be a record.

[23] IEEE. Standard Glossary of Software Engineering Terminology. In *IEEE Software Engineering Standards Collection*. IEEE, 1990. Std 610.12-190.

[24] See: http://www.interpares.org/display_file.cfm?doc = besser_eci.pdf.

real-time involvement of an officer or employee of the bank.[25] Such systems also receive records sent by other systems, for example when an ATM owned by one institution records a transaction against an account in another institution. In such cases, the system controlling the ATM sends a record of the transaction to the system of the institution which holds the account and, again, the transaction is completed and recorded without involvement of any human agent of either institution.[26] The receiving system subsequently completes other actions automatically, such as adjusting account balances and producing reports or other records accounting for such transactions.

In light of these examples, the definition of an interactive system needs to be expanded to "one in which each user entry *or input from another system* causes a response from or an action by the system."

As with interactive systems, for experiential systems, the InterPARES team looked for a definition that would provide a point of reference for selecting case studies. It decided to use Clifford Lynch's description of an experiential digital object as an object whose essence goes beyond the bits that constitute it to incorporate the behavior of the rendering system, or at least the interaction between the object and the rendering system.[27] This definition has not proved fruitful in qualifying either computer systems or documents created in them as experiential or not. InterPARES researchers have also conceptualized experiential systems as ones that immerse the user in a sensory experience. This concept, however, refers to a subset of the experiential computing described by Ramesh Jain, who, in addition to subjective experiences, such as story telling, folk computing, and personal event experience, identifies experiential applications used for administrative or research purposes, such as business-activity monitoring, homeland security, or bioinformatics. Jain depicts experiential computing as enriching cognition through sensation:

---

[25] Of course, bank officers/employees are involved in producing the record of the transaction in so far as they are responsible for establishing and enforcing the bank's ATM policies and procedures (which would make them the *de facto* writers of the record), and ensuring that these policies and procedures are translated into a system that in turn outputs appropriate transaction records in response to user requests/actions.

[26] Again, there is human involvement via consideration of who the competent writers and authors of the records are, as well as who the addressees are, all of whom are persons, not computers.

[27] Clifford Lynch. "Authenticity and Integrity in the Digital Environment: An Exploratory Analysis on the Central Role of Trust." In *Authenticity in a Digital Environment* (Washington, D.C.: CLIR, 2000), pp. 32–50, available at http://www.clir.org/pubs/reports/pub92/pub92.pdf.

> "[Users] must be able to explore and experience events from multiple perspectives and revisit them as often as needed to obtain that insight. In an experiential computing environment, users apply their senses directly, observing event-related data and information of interest. Moreover, users explore the data by following their own personal interests within the context of an event."

> "Experiential environments free users from the tedium of managing enormous volumes of disparate heterogeneous data. They don't try to interpret an experience; instead, they provide an environment that can be used to naturally understand events...."[28]

In this conception, experiential environments are necessarily interactive, but provide user interaction driven not by pre-programmed options, but by the user's interests, and they are likely to offer a greater variety of ways in which users can interact with the system than in interactive systems which are not also experiential. This conception promises to be more productive than the definitions adopted by InterPARES up to this point.

As regards dynamic environments, the InterPARES team initially defined them as "dependent upon data that might have variable instantiations and be held in databases and spreadsheets," thereby adopting Seamus Ross's view of dynamic computing.[29] However, in the course of research for this article, it emerged that dynamic is an attribute attached to a variety of environments.

> "'Dynamic computing' is a term that is used to describe flexible and adaptable approaches to tailoring computing resources to demands. It includes real-time distributed computing, adaptive computing, which can automatically adapt or configure computing resources to suit different tasks, and agent-based computing. Agent-based computing uses freestanding software agents that are variously described as intelligent, intentional, mobile, and remote acting. Such agents support applications, ranging from dynamic routing of network traffic, to automated email answering, business process management, military applications, and multidisciplinary problem solving in scientific research. Dynamic

[28] Ramesh Jain. "Experiential Computing," *Communications of the ACM* **46**(7) (2003): 48–54.

[29] Seamus Ross. 2000. *Changing Trains at Wigan: Digital Preservation and the Future Scholarship*. London: NPO Preservation Guidance Occasional Papers.

> computing has emerged as a dominant trend in the information technology in the twenty-first century."[30]

Such applications may employ a variety of techniques known collectively as evolutionary computing.[31] Dynamic systems are used in a variety of practical applications, such as engine testing in the aerospace industry and manufacturing execution system applications in the auto industry.[32] In many organizations, various categories of actors require different overviews derived from heterogeneous information and tailored to their different needs:

> "Business analysts would like to run decision support queries over the state and history of the company. Sales staff would like to understand the total relationship between the company and a given customer, including all past interactions and the current state. Financial planners would like to integrate overall budgets based on budget projections of units below them and would like to track previous forecasts with the actual operations of the business."[33]

Dynamic systems supporting such diverse needs are interactive and could be experiential. The information they present to users or to other systems is highly variable and contingent on multiple and diverse inputs from both users and other systems.

An interesting dynamic environment is that of systems which mediate interactions between organizations. Adaptability of such systems is seen as key to enabling organizations to form virtual alliances with loosely coupled business processes. A 'virtual alliance' is formed when interacting systems conduct actions or execute transactions as if the organizations had agreed to such collaboration beforehand. One of the values of dynamic systems is that, to enable such actions, they do not require that the organizations either integrate their respective business processes, or design either system to interact with the other.

---

[30] Summit Strategies. The 2005 Summit Seven: Dynamic Computing Gets Down to Business. Market Strategy Report 4EC-07. December 2004, available at http://www.summitstrat.com/store/4ec07detail.

[31] A. E. Eiben and J. E. Smith. *Introduction to Evolutionary Computing* (Berlin: Springer, 2003).

[32] Janos Sztipanovits, Gabor Karsai and Ted Bapty. "Self-adaptive software for signal processing: Evolving systems in changing environments without growing pains," *Communications of the ACM* **41** (5) (1998): 66.

[33] Narinder Singh. "Unifying heterogeneous information models: Semantic tags support knowledge webs," *Communications of the ACM* **41**(5) (1998): 37.

"The utility of complex systems is enhanced if the system can learn from experience and adapt its behavior. The ability of the system to develop and act on internal models that simplify the external world is basic to this mechanism. It allows the system to infer the results of actions before they are taken, and to choose actions that have productive results.... Next-generation systems must allow autonomous business object components to decide with whom to collaborate, what services to offer, what services to request, and what behaviors to exhibit."[34]

Subsystems of these systems, called "business object components," act as agents in the conduct of business. They may be programmed to make decisions and take actions; although such actions remain the responsibility of an official or employee, he or she does not carry them out personally. Such systems are clearly interactive, although the interaction may be limited to interactions between or among systems, without any direct, real-time, human involvement.

'Interactive,' 'experiential,' and 'dynamic' are thus attributes of systems, rather than types of systems. A system may be simply interactive, but an experiential or dynamic system is also interactive, and a system may be both experiential and dynamic.[35]

Interactive, Experiential and Dynamic Documents

The fact that a system is interactive, experiential or dynamic does not entail that documents made or received in it be themselves interactive, experiential or dynamic. Thus, we need to examine the information objects that exist in interactive, experiential or dynamic systems to

[34] Jeff Sutherland and Narinder Singh. "Application integration and complex adaptive systems Association for Computing Machinery," *Communications of the ACM* **45**(10) (2002): 59–64.

[35] The distinction among these three systems/attributes remains a matter of debate among InterPARES 2 researchers, primarily because they are more arbitrarily and subjectively than logically and objectively defined. As noted earlier, experiential and dynamic systems are types of interactive systems. However, in light of the need to compartmentalize them, another way of distinguishing them is by their relative levels of direct, real-time, human user involvement; with experiential and dynamic systems situated at opposite ends of the spectrum, and interactive system falling somewhere in between. In fact, from the definitions here provided, it seems that the only real difference between the interactive and experiential attributes is that only the latter ipso facto requires direct, real-time, human user involvement. One cannot help wonder whether the nuanced, largely subjective, distinctions among these three types of system attributes are: (a) actual and workable, (b) necessary, or (c) ultimately helpful.

determine whether any of them are documents. The InterPARES definition of 'document' as recorded information is a variation on the traditional diplomatic definition whereby a document is information affixed to a medium. This difference has important consequences. A "hard copy" document, by virtue of being affixed to a medium, has a fixed form and unchangeable content. But, in the digital environment, it is possible to generate something that to all appearances is a document, but is not affixed to a medium. Interactive, experiential or dynamic systems may display or otherwise present information that appears to be a document, but the system may not contain any object that corresponds exactly to the apparent document.[36] We might describe things that appear to be documents, but are not stored, as pseudo-documents or 'pseudocs.'

Live interactions, experiences, or dynamic processes do not necessarily produce or involve documents, even as a matter of appearances. For example, a dynamic system that monitors traffic on a network may simply trigger changes in the routing of messages to balance load across the network without producing any document about such actions.

A priori, there is no restriction on the form or content of a document, but it must be a finite entity: it must be possible to determine what information is and is not contained in that document, to specify its form, and to show how the content in that form constitutes an indivisible whole. With a hard copy document, the content, form, and wholeness of the document are embodied and manifest in the physical inscription on a medium. In the digital environment, the physical inscription of bits on digital media cannot reliably indicate what, if any, documents are written on those media.

One of the most important findings of InterPARES 1 was the recognition and articulation of the difference between the form in which an electronic document is manifested to a person and the form in which it is stored digitally. This difference is fundamental in two respects. First, it distinguishes a digital document from a traditional one, where the document is exactly what is inscribed on a physical medium in the way it is inscribed. Second, it makes necessary to describe the exact nature of a document and to determine whether it continues to exist across changes in the way it is inscribed on a digital

[36] As established in the first phase of InterPARES, a computer system may be said to contain an electronic document when it has the capability of reproducing that document. It may, but does not necessarily, do this by storing the document as a single data object.

medium. Without this fundamental distinction, we would not be able to assert, for example, that a document preserves its identity even when it is moved from a magnetic to an optical disc, or when it is translated from a word processing format to HTML for publication on a website.

The content, form, and wholeness of electronic documents are determined conceptually and logically rather than physically. A person's conception of a digital document depends on how it is manifested to him or her. It may be manifested on a screen or on some other output device. This manifestation is fundamentally different from the way the document is encoded and inscribed on a durable digital medium. The digital encoding, which is typically described by technologists in a logical model, enables a computer to produce or reproduce the intended manifestation, but it does not have the same form and in practically all cases will not have the same content as the manifested document. For example, a manifested document may be a textual narrative. It may be encoded either in character mode, such as in a word processing format, or as a document image, but neither the numeric byte values that correspond to printable characters nor the bits that are projected as pixels in an image have the same extrinsic form as in the manifested document. Such differences extend to other presentation features, such as organization into paragraphs and page layout. The content of the digitally encoded document will also vary from that of the manifested document because it includes data that indicate how to manifest the document. Simple examples are data that indicate presentation features such as line spacing, page breaks, and italics. More complex examples are specifications for extracting data from different tables in a large database, combining them with invariant data and presenting them as a single page form. There are many elements of the content of the digital components of a document that are not manifested to a person. If the manifested document is adequate to communicate the information intended by its author, the invisible or imperceptible digital elements may be necessary to manifest the document, but they cannot be said to be parts of the manifested document itself.

It is useful to distinguish three different types of data in which a document is encoded digitally: content data, which constitute the content of a document; form data, which enable the system to reproduce the document in correct form; and composition data, which tell the system what form and content data belong to which document. Form and composition data together determine the structure of a digital document; however, they are not equivalent to structure. The term 'structure' is commonly used in reference to the manifest organization of a document. Form data impose that organization on the content a document, while composition data tell the system what objects in storage need to be brought together to constitute the document and map them to the different elements of structure defined by the form data.[37] The three types of data – content, form, and composition – may be contained in a single digital component or separated in different components. For example, software recognizing that a stored object is in a word processing format will assume as a default that the object represents a single document, and contains all of its content and form data. But, if the same document were encoded in Tagged Image File Format (TIFF), each page of the document might be stored in a separate file. In that case, the system would use metadata about each file to determine which files belong together as a single document and in what order. Even in the case of a word processing document stored in a single file, some of the data necessary to display the document with the correct specific presentation features will not be stored in that file. Specifically, the data necessary to interpret the binary digits representing each letter or number for display in the appropriate font are stored in a separate dynamic load library file. In other cases, the form, content, and composition data may be stored separately. For example, in a database application, for the digital equivalent of a printed report, the specifications for the documentary form are stored in a report file, which does not include any content data, while the content data are stored in database tables, but it is most likely that the report only uses a subset of the content data and that it arranges it differently than the database itself. Composition data that map the data elements to be included in the report to the logical data model of the database are stored as an object called a database 'view.' To produce a specific instance of the report, additional composition data are needed, and are often supplied by the user who requests the report. For example, the report file for a monthly report of expenditures specifies the content and form for all months, while a user must specify a particular month.

[37] Archival literature frequently describes records as consisting of content, context and structure. The discussion above describes the relationship of the three types of data to content and structure, but not to context. This is because, as described in section 1 'Findings of InterPARES 1' above, the significant context is external to the record. It is constituted by the record's relationship to other records, the administrative environment in which the record is created and maintained, the action in which it is involved, and the persons involved in its creation. Some of the content of a record may indicate or reveal its context; nevertheless, it is made of content data.

Interactive, experiential and dynamic environments can produce the digital equivalents of traditional documents. When an interactive online catalog is used to transact sales, it should produce the types of documents needed in any sales system: orders, packing slips, invoices, receipts, etc. It does not matter whether such systems are enhanced with experiential or dynamic characteristics. Likewise, dynamic systems used to collect scientific observations should satisfy the requirements of the relevant scientific disciplines for reliable observational documents. Similarly, e-government systems used to obtain permits or licenses or to pay fees or fines should produce and keep the kinds of documents needed for such transactions, regardless of the characteristics of the system. For example, the Alsace-Moselle Land Registry, an InterPARES 2 case study in the government focus, is an interactive system used to make and receive electronic records in traditional documentary forms. The Registry produces electronic records which correspond exactly to the ordinance of inscription, the inscription in the register, and to associated records, such as contracts and cadastres, which have been produced in real property transactions for centuries.[38]

The situation is more complex when the objects produced by or contained in interactive, experiential and dynamic systems either differ significantly from traditional documents or have no traditional counterparts. The following sub-sections will describe interactive, experiential and dynamic documents separately; however, the discussion will illustrate that for documents, as for systems, these attributes are not mutually exclusive.

Interactive documents

An interactive object could be described as one which, when presented to a person or another system, allows the person or other system to input data that engender changes in the subsequent presentation of the same object. In a basic sense, all digital documents could be described as interactive because user interaction is required to select the document to be manifested, but this would be a trivial way of looking at interactivity. This analysis excludes interactions that are generic possibilities offered by the computer and

[38] Jean-Francois Blanchette, Francoise Banat-Berger, and Genevieve Shepherd, Computerization of Alsace-Moselle's Land Registry, InterPARES Case Study CS18. 21 September 2004. See also: http://www.interpares.org/display_file.cfm?doc=ip2_alsace_characterization.pdf.

not specific to a particular document. Generic interactions include selection of documents for retrieval and output, variation in the size of the window in which a document is viewed, magnification, viewing of one or more pages within a window, and accessibility features such as ones that change the size of text or render text aurally rather than visually. A document is appropriately described as being itself interactive only when it includes specific features that provide for user input and that use such input to change the content or form of the manifested document. The difference between generic and document-specific interactivity is illustrated in the options for navigating within a document. Navigation tools such as "Page Up" and arrow keys or a "Go To" box are generic options, while a hyperlink which enables a user to move from text that is currently displayed to another location in a document or to display content that is not stored as part of the document, is a document-specific option. A simple, but not primitive, interactive object consists of one or more sets of fixed data and related instructions (software) for selecting and presenting those data. Data input by a user could trigger specific instructions which select particular stored data and present them to the user. The user's input could also trigger other instructions that determine in what form and sequence stored data are presented. Examples of such interactive objects include web pages delivering government services online, musical performances based on human–computer interaction, and commercial video games.

Interactive documents might appear similar to traditional forms of documents, but their appearance does not reflect their substance, because it is limited to what the system presents at a particular moment. Assume, for example, that what the system displays looks like a document that could be printed on paper, as in the case of online sales systems, where stored content data include data that constitute the catalog of goods offered for sale, other data about which items are in stock, additional data about shipping and payment options, and data about individual orders, customers, and payments. A customer may browse the catalog, starting by selecting a category of goods from a textual list. The system would then display images and basic information about the goods in that category and it might allow the user to request a different image, a larger image, more textual information about a particular item, or reviews by other customers who had purchased that item. The user inputs are compositional data. The system changes what it displays in response to each user input. The form data are stored in one or more HTML files which specify how a web browser should display the selected content data.

The first difficulty with identifying such interactive objects as clearly defined documents is that the system does not store any object equivalent to what the user sees. Rather, it stores one or more databases from which content is selected, and one or more sets of instructions that interpret user inputs to select, retrieve and present some of the content. This difficulty, however, can be resolved by applying the distinction made in the first InterPARES project between the storage format and the representation format of electronic records. In interactive documents, any given presentation is a transient manifestation of a palette of possibilities provided in the stored digital components. The Preservation Task Force report of the first phase of the InterPARES project noted, "(r)eproducing an electronic record entails (1) reconstituting it, that is reassembling its digital components if it has more than one, or extracting any digital component stored in a physical file that contains more than one such component; and (2) presenting it in proper form."[39] It further distinguished digital components from the technological methods used to reproduce the records.[40] This distinction suggests that the digital components consist only of content data; however, the analysis of interactive objects leads us to clarify that the domain of digital components includes the instructions which select and present content in a given form (i.e., form data), as well as composition data which further define the content selected.

The second difficulty with identifying interactive objects as documents is that user feedback can change both the content and the form in which information is displayed. Unless the system keeps an audit trail of the user's feedback, the system literally cannot reproduce what the user saw.[41] Every instantiation could be different. However, the

[39] *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, available at http://www.interpares.org/book/interpares_book_f_part3.pdf, p. 6.

[40] Ibid. pp. 7–8. For a fuller discussion, see How to Preserve Electronic Records, Appendix 6 in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, available at http://www.interpares.org/book/interpares_book_o_app06.pdf.

[41] If the system does keep an audit trail, the relevant concern is not necessarily whether the system can reproduce every sequence and consequence of every "input → computation → output" process, but rather whether the system can reproduce only those process sequences and/or consequences that correspond to what the user (or the author/system, depending on the perspective involved) identifies as the documents/records that correspond to the user's interaction with the system.

system has the ability to reproduce the entire catalog or any selection of its content, and would produce the same instantiation in response to the same user inputs. Thus, a system can be said to keep an interactive document, regardless of any variability in form and/or content, when the system retains the ability to present that document on demand as a response to identical inputs. The fact that no user may ever see the entire online catalog is no more problematic than the fact that people only open a dictionary to look up one or a few words.

Experiential documents

An experience is a person's live, subjective involvement in or reaction to some event, activity, or entity. An experiential object is one which gives rise to experience or in some way captures an experience. Examples of experiential digital objects include works of electronic art, audio and moving images embedded in a web page, business monitoring applications which enable users to tap a rich variety of data sources, and virtual reality systems.

Experiential systems may produce or contain digital objects which are the electronic equivalents of traditional types of documents; for example, the script for a play or an audio recording of a musical performance can be either analog or digital. While computers create possibilities not available to an artist working in traditional media, the same could be said of different traditional media, such as oil and watercolor in visual art. Many works of visual computer art basically differ from traditional paintings and drawings mainly in the fact that they are digital.

Digital objects which include heterogeneous types of data, such as sound and moving images embedded in a web page, may appear more challenging to identify as unified entities, but are comparable to analog recordings of sound and motion video, which often are linked to related textual documents. As with visual art, there may be signficant differences in documentary form, due to fact that computers enable a greater variety of forms. For example, traditionally, the script and the film of a motion picture, and still photos and posters used for publicity, are different documentary forms, but in the digital environment they might be brought together on a web page, which can be treated as a single document. The heterogeneity of data types and the multiplicity of objects which may be combined in a single digital document in themselves pose no greater difficulty to identification and preservation than does a textual document on paper, which includes other data types, such as photographs or charts, or consists of several entities which could be independent documents; for example, a report with multiple attachments.

Excluding from consideration the live, personal involvement, two types of experiential digital documents can be distinguished: those used in producing an experience and those which capture it or, more specifically, those which record either the presentation which gave rise to a subjective experience or some aspect of the experience, such as the reaction of participants. A musical score embodies the first type of experiential document, one that enables a potentially unlimited number of performances by providing instructions for performing a work, while an audio recording of a performance, which captures a specific execution of those instructions, embodies the other type.[42]

The distinction between objects that enable performance and those that capture it might not seem to apply to the visual arts. Traditionally, in the visual arts, such as painting and drawing, artists produce works on physical media that can be experienced for as long as the physical instantiation endures. In the digital environment, works of visual art cannot be kept in the form in which they are accessible to humans, but only in some binary representation, which is independent of any physical medium. To experience the digital work, viewers rely on a system capable of reproducing the work from its digital components. Digital documents that enable reproduction of static works of visual art are comparable to recordings of performances in the performing arts, as they both freeze an end product in the creation of an artwork and enable that specific instantiation to be retrieved subsequently. However, digital documents that enable presentation of

[42] There is some disagreement among music theorists on whether the score is also a musical work in its own right, but this does not invalidate the the ideas presented here.

interactive visual art belong to the class of objects that enable performance.[43]

Pioneering computer artist, Myron Krueger, for example, has created computer mediated installations which respond to viewers' inputs received from a variety of devices, project computer-adapted video of the viewers, and enable audience members to alter the video projection by virtually "touching" the projected images of themselves, other members of the audience, or computer-generated objects, ranging from graphic images drawn by the artist in real time, software-generated strings and ovals to animated organisms.[44] The content of these artworks consists of data produced during experience of the work: visual objects created algorithmically from data input from various sensors, which detect the presence and movements of spectators, and other visual objects either drawn by the artist observing the spectators' movements in real time or generated by stored algorithms but modified based on data from spectators. In a basic sense, such works have no fixed content, but viewed at a level of abstraction one step removed from that of the work as experienced, the content is bounded by the possibilities allowed by the artist either in the sensors and

[43] The digital environment enables an artist to record an artwork in different documentary forms. In addition to the documents enabling performance and the recording of a performance, the InterPARES 2 case study Waking Dream is documented on a website, which is regarded by the principal author as part of the art work, and which includes the 'script' along with information about the performance space and the gadgets used, résumés of the artists, and videos and images from performances. The web page is a document whose digital components include the web page itself and the objects accessible from the page, such as documentation in textual form, audio-visual recordings of actual performances, and sample still pictures of imagery projected during performance. The computer code is neither included nor described on the website. The web page is neither a recording of a performance of Waking Dream, nor a document enabling a performance, such as the script of a play or score for a musical work. Rather, it is an alternate form of presentation of the work. For a complete description of the case study, see: http://www.interpares.org/display_file.cfm?doc=ip2_waking_dream (complete).pdf. The Waking Dream web page illustrates a third type of experiential document, one that enables a user to experience a work by interacting with the system: the nature of the interaction is in fact more active than passive, or more participatory than observational. This is a more complex form of an interactive system, possibly including heterogeneous types of data and correspondingly more varied possibilities for presenting those data, and providing more flexibility or sensitivity in responding to user input.

[44] Söke Dinkla. The History of the Interface in Interactive Art. 1994: http://www.kenfeingold.com/dinkla_history.html. See also: http://a.parsons.edu/~praveen/thesis/html/wk05_1.htm1, and http://www.artmuseum.net/w2vr/timeline/Krueger.html.

projectors (input and output devices), which form each installation, or in the computer programs the artist wrote for the work. The computer programs, as well as the documents describing the installations, fall into the category of documents that enable performance. While users' actions in marketing applications indirectly influence the selection and form of presentation of content data, in interactive art installations user actions not only directly shape the performance, but also provide part of the content.

From such examples and case studies of visual art, InterPARES is beginning to advance the proposition that, in the digital world, the difference between performing arts and other forms of art is disappearing, in that artists in any field can produce digital works that can only be manifested over time by re-creating them on the basis of a set of instructions and related information needed to carry out the instructions as intended by the artist.

The InterPARES case studies of electronic music indicate that the set of instructions recorded by the composer – which might include a score, computer codes and other instructions on performance – may not be sufficient to reproduce the piece: the work may also require specific software patches, hardware or other devices, such as a synthesizer, and even a specific kind of interaction between the performer(s) and all of the above. Such interaction so far has never been described in a way that can be reproduced. On the basis of the case studies results,[45] increasingly, both composers and InterPARES researchers are arriving at the conclusion that a work of digital music can only be reproduced if the author describes each digital, intellectual and performing component of it and the interactions among them, by producing a set of instructions for re-creating each part of the piece and the piece as a whole. Thus, the case studies of digital music, as well as digital theatre – as we will see below with the case study Waking Dream – reveal that the class of experiential documents which enable experience includes at least two subtypes: one consisting of instructions whose execution produces a performance, and the other describing the components, context, preconditions, or requirements for performance whose execution allows for future performances. In other words, the artist will have to become an active participant in preservation.

[45] For example, "Obsessed again..." an interactive piece for bassoon and computer by Canadian composer Keith Hamel. See http://www.interpares.org/display_file.cfm?doc = ip2_obsessed_again(complete).pdf.

The distinction between experiential documents that enable performance and those that capture it is valid for non-traditional art forms as well. Waking Dream, for example, is a multimedia theatre artwork that explores the border between being awake and dreaming. In performance, two dancers personifying 'Wake' and 'Dream' move about the stage and through the audience, accompanied by a soundtrack, while both live and digitally recorded videos are alternately projected. One of the two performers wears a head-mounted camera and holds a remote control device, which allows or blocks the video projection.[46]

The digital objects that enable performance of Waking Dream include: a textual document that describes the performance and provides details for staging it, a collection of core sound samples that is remixed prior to each performance to create its sound track, digitally recorded images, and a specially developed computer code, through which the computer mediates interactions among the camera, remote control device, video projector, and audio devices. Like an interactive sales catalog, the digital objects that enable performances of Waking Dream include content data – audio and video in this case – and instructions for presenting the work, which constitute both form and composition data. Similarly, an instantiation of the work involves variable sequencing of the stored content data, and the presentation of the data depends on the interactions from one of the dancers. All four types of objects might be considered as documents in their own right, since they each have fixed form and content. The second type of experiential digital document also exists in this case. A performance of Waking Dream has been recorded in Apple Quicktime '.mov' format: this file reproduces that performance, or at least those aspects of the performance that were capable of being recorded, since much of the performance happens almost in the dark, with only infrared light.

A very different type of experiential object is created in online marketing applications. Unlike artistic works, which enable unique, subjective experiences, online marketing applications aim at producing an experience that will lead to a specific behavior; namely, the purchase of the goods or services offered by the sponsor of the application. Superficially, such sites appear comparable to online sales applications, but unlike online sales catalogs, marketing applications may not even offer the possibility of online transactions. For example, websites sponsored by pharmaceutical companies do not offer for sale medicines requiring a prescription, but such sites have proven very effective in building brand loyalty and inducing their visitors to talk

[46] InterPARES 2, Waking Dream Case Study Final Report, cit.

to their doctors about a particular medicine. These sites are designed in recognition of the facts that most visitors to a medical site will not return and are reluctant to provide personal medical data online. Given such barriers, it would seem unlikely that these websites could significantly influence behaviors, but they do. Rather than providing a visitor with information that directly supports a sale transaction, marketing sites are designed as personalized information resources. They do not ask visitors to input information about themselves, but minutely observe online behavior, gathering information from visitors' actions on a site to anticipate their individual requests for information and respond to them with the most appropriate marketing message: "Conscious of it or not, even the most tight-lipped visitors communicate to every website they visit. Every keystroke is a clue to their situation, needs, and preferences – if the site can interpret and act on it."[47] The digital components of such websites include HTML documents that provide the basic documentary form of the site, a store of relatively atomic messages that can be delivered in response to user inputs, rules for collecting data about users' navigation of the site, rules for interpreting such data as the basis for selecting and presenting stored information, and possibly rules about saving user inputs to create profiles either of individual users or classes of users. Such profiles are subsequently used to refine determinations about what messages to present to users and in what form. In contrast to an online catalog, whose content is likely to be stable at least for some finite time, sites which continuously collect data about user interactions with the site and use the data to modify subsequent presentation of content create documents which, in effect, are always in progress. They are never finished, unless the application is terminated.[48]

A visit to a marketing website can be regarded as an experience and the information provided by the site during a visit can be considered as analogous to a performance, though it might be better termed a production. Accordingly, the stored messages, the HTML documents, and the rules can be said to enable the production, and the data captured about user navigation of the site can be regarded as

---

[47] David Reim. "Online behavior: A brand builder's best friend," *Pharmaceutical Executive* **22**(4) (2002): 104–108.

[48] In this and many similar cases, the most likely and workable solution to the absence of a clearly identifiable entity that can be kept as a document is to establish intermediary 'termination' points (these could be predefined or randomly chosen to facilitate statistical analysis of the results, if desired, or they could be triggered whenever a predetermined set of user interaction criteria are met, etc.) in the ongoing process, when documents are produced that attest to the state of the system at those points.

composition data that subsequently determine both the content and form of information presented on the site.[49] The user navigation data might also be considered as a recording of an experience of that site, but such data are not used to reproduce the experience, and are unlikely to be organized in a way that would make it possible to reproduce the experience.[50]

In general, documents that enable performances also enable variations in performance. Performances vary, depending on how much discretion the instructions give to the system executing them or the performers interpreting the work, on the ability of the performer(s), the characteristics of the instruments or devices used, the characteristics of the performance spaces, etc. As with purely interactive documents, the documentation of an artwork may call for variations in content; for example, the remixing of core sound samples, the alternation between live and recorded video, and the differences due to movements of the dancers in Waking Dream. While there is an element of artistic license whenever live performers execute written instructions, and the limits of artistic license are subject to argument, there are also, inevitably, boundaries to the variability. The narrative describing the movements of the dancers in Waking Dream, for example, is set out in broad terms, but the work must be performed by exactly two dancers, one of whom controls the video projection. If the core sound samples or the stored images were changed, the result would be a different artwork, albeit one closely related to the first. Variability in performance based on documented instructions is common to both digital artworks and those recorded in more traditional forms.

A specific performance captured in one or more documents has narrower bounds of variability than that permitted by the documents that enable performance.[51] Nonetheless, the reproduction of a recorded performance will vary depending on the quality of the recording and the system used to reproduce it. While some works of performance art may require elaborate and complex arrangements for

[49] Of course, one should not be induced to believe that the content and form are unique entities created on the fly from an infinite pool of possibilities, as they are simply manifestations of predetermined combinations of content and form that are then selected from a finite pool of such combinations in response to user behavior.

[50] To a degree, this is precisely what the history cache in a web browser does, which users are able to reproduce by using the browser's back and forward buttons.

[51] Because the documents in which the performance is captured are 'static' and/or are incapable of capturing all aspects of the actual performance, they ipso facto must embody less variability than is permitted by the rules describing the components, context, preconditions, or other requirements for the performance.

performance, in the digital environment, the reproduction of a performance may be accomplished simply by the reproduction of the document in which the performance is recorded. For example, the presentation of music or other sound recorded in digital audio format, such as '.MP3,' is accomplished simply by "playing" the file on a computer with an appropriate sound card, software and speakers. Similarly, regardless of the complexity of the software used to create them, static digital images of visual art, as well as many cases of motion video, can be reproduced simply by presenting the digital document in which they are recorded.

Reproduction of a digital document is not sufficient to reproduce a performance; for example, in cases where there are special requirements, such as that a person experiencing a performance be in a specially designed space, or where the performance requires special gadgets, such as virtual reality goggles. Clearly, both the format(s) of the document's digital components and the system used to render a recorded performance must be adequate for the work. This can range from the simple case where digitally recorded music cannot be reproduced on a computer which lacks either a sound card or speakers, to complex ones where the person experiencing the performance must have appropriate means of interacting with the system used to reproduce it. But once this threshold is achieved, the question is one of the qualities of the reproduction, not whether the work is reproduced.

In sum, in experiential environments we have found two types of documents which may themselves be qualified as experiential: (1) documents that enable performance or production of a work, including both the documents which describe the work and/or the instruments, devices or other things used in the performance of the work, and those which provide instructions on how to perform the work, and (2) documents that capture a specific performance or experience. The essential purpose of characterizing a document as experiential is to emphasize that it is clearly intended not merely – perhaps not at all – to communicate specific information, but to engender subjective experience.[52]

Dynamic documents

As with interactive and experiential systems, dynamic systems may produce digital equivalents of traditional documents. Dynamic

[52] This is consistent with the definition of 'experiential' by InterPARES researchers, but broader. The project's definition is restricted to 'sensory experience,' but an experience could be intellectual or affective, as well as sensory.

systems may, for example, interact with other systems to conduct business transactions or even form contracts without human mediation, but the documents which are produced in such interactions should conform to requirements for records of transactions and contracts.

However, dynamic systems also produce objects which could be described as dynamic themselves. An object may be said to be dynamic when (i) it has a fixed form, but draws its content in real time from other sources, (ii) the content data available for presentation are fixed, but their presentation, both in form and selection of elements of content to be presented, varies in response to real-time inputs from a person, another system, or an input device, or (iii) the content data, though stored as part of one or more digital components, change frequently by additions, deletions or replacements. These types of dynamic objects are not mutually exclusive. An object may belong to all three types; however, belonging to any one of them is sufficient to categorize it as dynamic. Moreover, all three types are also interactive and may be experiential.

An object that dynamically acquires content data may have a fixed form in that its extrinsic and intrinsic elements manifest themselves in such a way that the appearance of the document and its intellectual structure do not change, as in the case of websites which present information on the weather or on international currency exchange rates. Such objects can be described as documents that have fixed form, some fixed data and some variable data. Fixed data on a weather site include, for example, the locations for which weather data are available and, on an exchange site, include the currencies for which exchange rates are given. The variable data include current temperature, precipitation, exchange rates, etc. Some digital documents of this type may also allow variations in the way content data are presented, similar to interactive documents whose content data are stable. In these cases, the variation in presentation is governed by rules or instructions.

This type of dynamic document can also be found in many other systems, including the marketing applications or interactive visual artworks described in the last section on 'Experiential Documents.' The source of the content data in dynamic documents of this type could be a scientific instrument in a laboratory, a satellite transmitting live imagery, a video camera aimed at a highway, or wireless equipment transmitting heart rate, blood pressure, or other biometric data about an outpatient, or any number of other sources external to the system.[53] This type of document may have some fixed content, only variable content, or a combination of the two. Objects which acquire, process and present, but do not keep, data from external sources are analogous to temporary or intermediary printed forms; for example, when ordering an item at a tools store outlet, one writes the item number and description onto a temporary paper form, hands it to the clerk who then enters the data from the paper form into a permanent order form on the company's computer system and throws the paper form away. Someone in the warehouse then views the order form on a computer screen, pulls the order from the shelves, then prints a hard copy of the computer version of the order form. The printing is necessary in that type of transaction because, while variable data entered on a paper form become part of the document, with online 'forms,' the data are included only in the display of the document, and elements of it are discarded or replaced in response to successive user inputs or other external stimuli. Because the data from external sources are not kept within the system, they are part of a document only while it is being presented by the system or, of course, when the form is printed out. As with other cases we have considered, the form and content of a document are determined by instructions which govern processing of the external data.

Documents in online marketing applications also belong to the second type of dynamic objects, where the content data available for presentation are fixed, but the selection of content data and the mode of their presentation vary dynamically. These include web applications that enable persons to explore a website using a variety of options. Such applications continuously collect data about user actions on a site, such as how many seconds a user spends on any portion of the site, where each mouse click occurs, etc., but such data are never presented. Rather, they are used by the application to determine what stored messages to display next and how to present them. Such documents have fixed total content, consisting of information about the site's sponsor and its products, continuously changing composition data about visitor interaction with the site, and a specific documentary form and content determined by the rules for interpreting data

[53] Nicola Ferrier, Simon Rowe, and Andres Blake. "Real-Time Traffic Monitoring," In *Proceedings of the 2nd IEEE Workshop on Applications of Computer Vision, Sarasota, Florida, 5–7 December 1994*, pp. 81–88, available at http://ieeexplore.ieee.org/iel2/998/7985/00341292.pdf. Noel Baisa. "Designing Wireless interfaces for patient monitoring equipment," *RF Design* April 2005: 46–54, available at http://www.rfdesign.com/mag/504rfdf4b.pdf.

about user interactions. Online sales applications that collect data about user interactions and use them to determine presentation also fall in this category. It might be noted that this is exactly the same process described above for experiential documents associated with marketing websites.

The third type of dynamic object, where stored content changes frequently, is actually a variation on the second, and is found in e-government applications. One of the InterPARES 2 case studies, VanMap, is an example of this type of dynamic object. VanMap is a GIS system that allows the City of Vancouver to meet the needs of city officials and employees in providing services to Vancouver's citizens and businesses. VanMap supports the functions and activities of the following departments: Community Services Group, Engineering Services, Corporate Services Group, Board of Parks and Recreation, Vancouver Police Department, Fire and Rescue Services. Decisions about how to organize information into GIS layers and what sets of data each layer should contain are made collectively by the departments and the VanMap Technical Team. Data are uploaded by each department directly in Oracle Spatial or taken as extracts from external offices databases (for example, permit and license data stored in PRISM or License+ are extracted to an SQL server; property tax data are extracted from the SQL Property Tax System, etc.) for inclusion in VanMap by the Technical Team, which is responsible for its administration. Engineering and constructive solid geometry (CSG) graphics are created in the form of CAD drawings in AutoDesk, or keyed or drawn in the Oracle Spatial database. VanMap data are overwritten at each update and, every once in a while, existing layers are modified to receive different kinds of data sets, and new layers are added. The data about transactions which the system supports are replaced or erased whenever any data used in support of a process are updated, or new data layers are added, or whenever the instructions for a process are modified.[54]

Dynamic systems often support extraction and processing of information from heterogeneous sources, where the sources themselves may vary even in the course of a single run. Dynamic systems also vary in the output they produce. Variations in the types of data in the input stream change both the possibilities and the requirements for processing input and, therefore, require systems that can reconfigure themselves on the fly. Changes in the way the system executes its

[54] For more information about the VabMap case study see: http://www.interpares.org/display_file.cfm?doc=ip2_vanmap_characterization.pdf.

processes may occur autonomously as the system assesses in real time the external data sources supplying data, what data are supplied, or the characteristics of those data. In the face of such variations in data, the system may invoke different software agents or components or, in more advanced cases, may modify the software it uses to present the data. Changes in software may in turn modify the content data, for example, by applying different calculations on raw data. Such techniques are used in practical applications such as scheduling and modeling of financial markets, as well as in computer art and "edutainment;" that is, applications which achieve educational goals through entertaining means.[55]

Conclusions on the Classes of Digital Documents

The preceding three sub-sections show that interactive documents constitute a major class of documents that can only exist in a digital environment. This class is distinguished from other digital documents that do not have document-specific features enabling user interactions that alter the form or content of the document when it is manifested. Thus, we may postulate a basic division of digital documents into static and interactive categories. The preceding sub-sections also show that dynamic documents are a subset of interactive ones, characterized by the fact that their variability derives, at least in part, from variation in the rules used to generate the document.[56] Experiential documents, however, do not constitute a separate category or sub-category within this scheme. Distinguished on the basis of their role in engendering subjective experience, experiential documents may be static, interactive, or dynamic. For example, a visual artist may create a static digital picture which is entirely analogous to a traditional painting. But the artist could also add interactive features to such a picture, and could use dynamic algorithms to change the way the image is generated, as well as its form and content.

This analysis enables us to construct a taxonomy of digital documents, as shown in Table I.

---

[55] Dale Thomas. "Aesthetic selection of morphogenetic art forms," *Kybernetes* **32**(1–2) (2003): 144–155.

[56] One might argue that dynamic documents are not necessarily a subset of interactive ones. It is possible to conceive of dynamic documents that do not include any interactive features but generate varying displays from algorithms that alter themselves. However, such cases would be properly characterized as pseudocs or even applications, rather than documents.

Table I. Taxonomy of Static, Interactive, and Dynamic Documents

Class	Description
1	Static documents Digital documents are static when they do not provide possibilities for changing their manifest content or form beyond opening, closing, and navigating within the document. Once a static document is retrieved and manifested, its entire content is available to the user and its structure is invariant. A user may need to interact with the system in order to access the content, or different portions of the content, but such interactions do not change the form or content of the document. Any user exercising an option for navigating within the document – which includes options for different manifestations of the document – will be presented with the same result.
1.1	The electronic equivalents or counterparts of traditional documents *Examples* *Letters; reports of scientific experiments or observations of natural phenomena output from dynamic systems; digital sound recordings, digital motion video, and visual art works*
1.2	Documents that have no exact counterpart in hard copy or analog form but have fixed documentary form and content[a] *Examples* *Snapshots of web pages, and recordings of performances of artworks which have characteristics that may exist only in a digital environment, as well as the results of freezing and capturing the output of a system that modifies its own instructions for processing or presenting content data*
2	Interactive Documents Documents that present variable content, form, or both whose rules governing the context and form of presentation may be either fixed or variable
2.1	Interactive Documents which are not dynamic Documents where the rules which govern the content and form of presentation *do not* vary, and where the content presented in any instance is selected from a fixed store of data within the system *Examples* *Online sales catalogs, interactive web pages, and documents which enable performance of music and other works of art*
2.2	Interactive Documents which *are* dynamic Documents where the rules which govern the content and form of presentation *may* vary
2.2.1	Documents where the content and/or its presentation vary because it includes or is otherwise impacted by data that change frequently

Table I. Continued

Class	Description	
	Examples	*Documents in systems whose design permits updating, replacement or alteration of data but does not provide for keeping older or superseded data, and websites that collect data from users or about user interactions with or actions on a web site and use those data either to generate or determine subsequent presentation*
2.2.2	Documents where the content varies because it includes data received from external sources and not stored within the system	
	Examples	*Websites which present information on topics such as the weather or currency exchange rates, as well as many interactive artworks*
2.2.3	Documents produced in dynamic computing applications, which select different sets of rules – software applets or service components – to produce the documents depending on variations in user inputs, in the sources of content data, and in the characteristics of that content	
2.2.4	Documents produced by adaptive or evolutionary computing applications, where the software which generates the documents can change autonomously	
	Examples	*Websites which involve the scheduling and modeling of financial markets, as well as some types of dyanmics computer art and 'edutainment' sites*

[a]This class includes the outputs of any method of capturing or freezing something presented by an interactive, experiential, or dynamic system.

Interactive, Experiential, and Dynamic Records

Can interactive, experiential or dynamic documents be records? As established in InterPARES 1, two of the essential characteristics of an electronic record are a fixed form and an unchangeable content. However, the constraints of fixed form and unchangeable content are not absolute. Records that have suffered some loss or corruption of elements of form or content through accident, mishandling or environmental factors remain records, provided the loss or corruption does not compromise their nature. Such alterations must be considered case by case.

For electronic documents in general, fixed form does not mean a completely invariant form, always identical to itself. A textual document on paper has an immutable form: the alphanumeric characters have a definite size and configuration; margins are firmly set, etc. But, as explained in distinguishing interactive from static documents, both the technology used to manifest digital documents in individual instances, and different choices made by users viewing them can cause variations in the manifest form and/or content, even when there is no variation in the stored digital data used to generate the manifested document.

'Record' comes from the Latin, *recordari*, to remember. The essential function of a record is to serve as a bridge over time, to carry information about an action, event, or state of affairs forward for when it is needed in subsequent actions or for reference about what happened or was described or said in the past. Setting aside variability due to generic features of digital technology, static digital documents clearly satisfy the requirements of fixed form and unchangeable content, regardless of the characteristics of the systems in which they are made or received. Examples include purchase orders and invoices created in transactions executed in online sales applications. In cases of interactive documents whose form or content varies according to fixed rules (class 2.1 in Table I) and documents whose form or content varies according to rules which may themselves be variable (Table I, class 2.2), the variability of form or content will prevent the documents from serving as records. A record that does not contain a fixed message or convey that message in a fixed form cannot be recalled and cannot be a means of remembering.

However, there are cases where the content or form may vary but in a way that does not prevent the documents from serving as records. In many interactive, experiential and dynamic documents, authors or writers[57] intentionally use specific possibilities which digital technology offers for variability in the form in which information is presented. In such cases, the form is 'fixed' in that the design allows certain aspects of form to vary and not others. Documentary forms that include variable elements do not violate the requirements for fixed form, any more than analog audio and motion video recordings, which present temporal variations in sound and imagery. Such variability in presentation intended by the author should be seen as part of the extrinsic elements of the documentary form. In digital documents in which fixed rules govern variation in content and/or

[57] 'Author' and 'writer' are terms used here as defined by diplomatics. See Footnote 6. When the author is an individual, it usually coincides with the writer. When the author is an organization or a collective or collegial entity, the writer is the person(s) who articulate(s) in writing its will, usually the signatory(ies).

form, (class 2.1 in Table I), such bounded variability is not a product of information technology in general, but is embodied in specific digital components of the document, such as interactive forms, software applets that generate varying presentation, business rules, software that uses user input to determine subsequent output, and rules which enable systems to adapt to changing inputs and demands. However, it is difficult to conceive how a dynamic document, wherein the rules which govern form or content change, could be a record, except perhaps as a 'draft' which is in the process of being developed for as long as it remains in the dynamic system.

With electronic records, then, the 'fixed' form consists of those aspects of form which the author or the writer intended or could control. While there may be difficulties in discerning a person's intent, generic variability enabled by information technology should not be considered as expressing an intention of an author or writer. For example, the author of a textual document probably intends, or at least expects, that a document will be displayed to readers with the same type size, line length, colors, etc. the author sees on the screen. But, in some situations, the author has no means of preventing variations of the type described in the last paragraph; therefore, the effects of different hardware, user selected window size, and other aspects of variability due to the technology used to view or experience a document after making it cannot be considered as intended by the author. Aspects of form that require a specific intentional action by the author or writer, such as division of text into sections, inclusion of images in a textual document, and any differentiation of the appearance of a portion of the content, either from adjacent content or from the norm for the rest of the document, convey the author's or writer's intent. There is also at least an element of intent in an author's or writer's choice of digital data type or format.

An author or writer may establish bounded variability in the content of a digital document. A significant class of documents where variability in content does not negate the necessity for fixed content is that of documents which allow variable subsets of the content to be displayed at any moment. Class 2.1 documents in Table I, illustrated by interactive sales catalogs, are reproduced through processes which include options enabling users to select content. Such options might be seen as analogous to the variation a user would see when browsing selectively or randomly through a catalog printed on paper. Although one might argue that the situation is not truly analogous to a selective reading of a printed document, because the possibilities for selecting digital content are not entirely, or even primarily, at the user's discretion, but depend on rules that are part of the document itself, these rules are not fundamentally different from the restrictions placed on the users of a printed catalogue by the catalogue's physical constructs and design 'rules' (e.g., layout, product categorization, pagination, etc.). In such cases, while what is presented to the user at any moment may appear to be a document, it is instead part of the only existing document, the catalog. The full content of an online sales catalog comprises all data stored and available for presentation to a user. It includes data about the goods offered for sale, such as textual descriptions and images, as well as data about related topics, such as payment and shipment options. The fixed form of the catalog includes those aspects that are always shown as well as those that determine how selected content is presented, such as the size and position of images, and whether the catalog exhibits various categories of data in windows separate from the main display. The digital objects that enable the selection of content are digital components of the document. This qualification does not negate the necessity of fixed content of the document as a whole. It merely recognizes that interactive digital environments enable an author or writer to structure a document so as to permit variable selection of content and variable sequencing of that selection. Cases where the documentary form permits selective display of subsets of the content can satisfy the requirement for fixed content.

A document that gathers some or all of its content data from external sources and does not store them concurrently – as opposed to sequentially – within its digital components cannot be presumed to contain a fixed message. However, in some cases even documents in this category can be records. Like a musical score or script for a play, a document that delineates a fixed form in which external data are to be presented and may include some unaltered content may be an instrumental or instructive record. Myron Krueger, for example, has created artworks where the software generates geometric objects, which are projected on a screen where they move and change their size and shape; however, the specific shape, size, movement, and the sequence of changes – the specific content – depend on characteristics or actions of individual viewers. In contrast to an online sales catalog, all of whose content is concurrently stored in a digital document, in such artworks the digital document that generates the presentation does not contain all of the content data, at least not cuncurrently. This document determines what objects may appear in the artwork, their basic characteristics – for example, it defines whether one object is a closed geometric loop; it makes another look like a small animal; and presents the silhouette of a spectator captured in real time as a third – and their possible behaviors – for example, it establishes that a "Critter" will seek to move to the highest point of a person's silhouette. The variable content which is manifested in a given instance may be captured in a recording of the presentation, but it is not stored in the document which enables the presentation.

There is a commonality between such artworks and online sales catalogs in that, in both contexts, the document that is kept and used for future reference is the digital object that is stored in the system. It is not the materialization of that object on a computer screen or other rendering mechanism. Provided that the other requirements for being a record are met, the record in such cases is the digital entity, not the human-perceivable form which is reproduced from it.

Applications may display pseudocs whose presentation includes both selected subsets of stored content and data obtained from external sources. For example, if a user of an online catalog inquires about the current availability of an item, the sales application may send a query to the inventory database which tracks items in the company's warehouse, or it may even query the inventory database of an independent supplier. The application would use the response from the inventory database to inform the user whether the item is currently available and how soon it could be shipped, displaying this information as if it were part of the catalog. Such external data are intentionally not stored as part of the catalog because they would quickly become outdated and misleading. Behind the variable content, the catalog document must include fixed rules that enable the system to get and present the variable data in real time. These rules are part of the stored digital document.

The description of performance art in the 'Experiential Documents' sub-section above led to the distinction between documents of performance and documents which enable performance. From the current discussion, we can see that this distinction applies to other domains. Interactive documents, with one exception, are "enabling" documents. They enable performance of artworks, execution of business transactions, or conduct of experiments or carrying out of programs for collection and analysis of observational data. The exception is represented by documents where changes in content data do not reflect an explicit intention of the author, but rather result from a failure to provide for the retention of data in the system and/or to compensate for system changes. This is illustrated in the case of VanMap, where data are regularly overwritten and the data model is changed on occasion. Such changes result in the inability to reproduce documents previously created in the system. This inability is due either to the fact that some of the content of such documents is no longer available or, even when all of the original content remains available, to the fact that changes in the data model can result in the inclusion in subsequent presentations of the document of new categories of data which were not present in the original version, or in the selection, processing or presentation of the original content in a different way.

This analysis leads to a conceptualization of an electronic record which is notably different than that articulated in the first phase of InterPARES. In InterPARES 1, an electronic record was any record manifested by a computer system to a human or another system. The form of the record is that of the document manifested by the correct processing of the stored digital components. The stored components enable reproduction of the record, but are not the record. This distinction between the manifest record and its digital components is basic and essential because errors in processing the stored data could result in failure to reproduce the record or in the production of a different document and because it is possible to preserve the manifest record – that is, to maintain the ability to reproduce it – even when the digital components are altered, for example, by reformatting or migration to different digital media. In this view, all of the essential properties of the record are found in the document that is manifested, and these properties are basically independent of how the document is encoded in digital components. Thus, there is an inverse dependency between the record and its digital components: the record is produced from its digital components, but the components must be produced in such a way as to guarantee that all essential properties of the record are present and identical whenever the record is manifested.

These findings were not wrong, but too limited. Given the essential memorial function of a record, the digital components might themselves constitute a record or a set of records, depending on how they are instantiated in the system. The digital components and the document reproduced from these components may constitute, that is, related but distinct records: the digitally stored record(s), and the "manifested record," which can be defined as the visualization or materialization of the record in a form suitable for presentation to a person or another system. The primary purpose of keeping the stored record is to be able to reproduce the manifest record, while the manifest record is preserved to communicate information to persons or other systems. The study of interactive records in InterPARES 2 further enriches the concept of the manifested record to encompass any and all variability of form and content which is specific to the document. InterPARES 2 case studies also lead to the recognition that a digitally stored record includes not only the data which must be processed in order to reproduce the manifest record, but also the rules for processing the data, including rules which enable variations in the content or form of the manifested record. Another reason for differentiating between the stored and manifested records is that one or more of the digital components of a manifested record may also be used in reproducing other manifested records. If a stored object is used to reproduce more than one manifested record, it cannot be the equivalent of any one of them. Moreover, if in a set of digital components which together are used to reproduce a manifested record, any one component is used in the output of multiple records independently of other members of the set, then that component should be considered a record.

The findings may seem radical, but in fact there are well established precedents. In medieval Europe, when the profession of notaries became so powerful that most transactions had to be recorded and preserved by them, they introduced efficiencies by not writing out the records of the transactions that they witnessed. Rather, they would take a parchment, fold one corner forward, and write on it the transaction type, the names of the parties, the date, the description of the transacted property or matter, and any other data specific to that transaction. Then, they would file away the blank parchment with the annotated corner, called *imbreviatura*.[58] At the end of each year, they would bind all the *imbreviaturae* of the year in a volume, and index the volume and/or keep a separate registration of the transactions bound in that volume in a book of *regesta*. If, later on, one or more of the parties to that transaction, or their descendants, wanted the complete record of the transaction, the notary would find by date the volume containing the *imbreviatura* in question, retrieve the document in it through the index or the register, take a new piece of parchment (or paper, if appropriate), and write out a complete record following formulas contained in a special book, called *formularium*, which provided clear instructions for writing out a record for each type of transaction that occurred in a specific range of years, and inserting the specific data written on the *imbreviatura* corner at the proper locations. Thus, what the notaries maintained was not the complete record of each transaction in its final form, but a record of the content of the transaction and another record of the documentary form in which that type of transaction had to be manifested. Rather than keeping transactional records per se, they maintained the ability to

[58] Sometimes, rather than on a corner, they would write the data on the back of the medium.

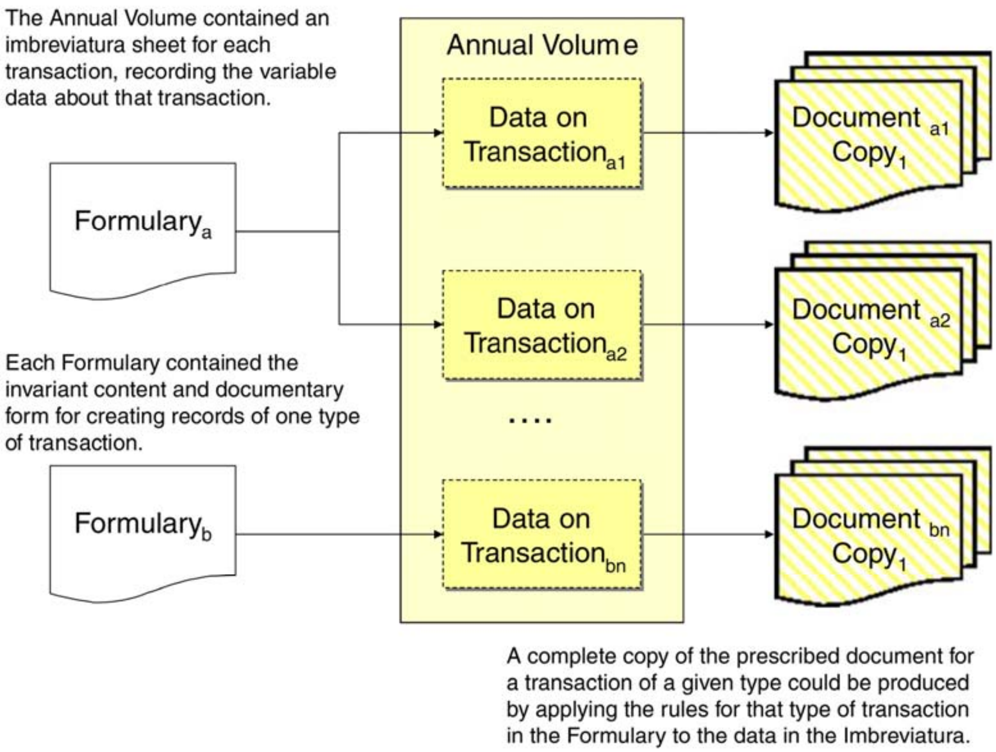


Figure 1. Keeping and Producing Records Using Formularies and Imbreviaturae.

produce an authentic copy of such a record upon request. In this system, each *imbreviatura, formularium, register, and reproduction of a transactional record is a record.*

The notaries kept a record of the fact that a transaction had occurred (register and/or index), a second record defining the documentary form for each type of transaction along with identification of the variable attributes needed in each type of records (*formularium*), and a third record containing the data values of those attributes for each transaction (*imbreviatura*). Combining the document model contained in the *formularium* with the specific values in the *imbreviatura*, they and their successors could produce the accurate and authentic record of the transaction when needed, even centuries later. Each record of the transaction produced in this manner would be an original record. However, because of the trustworthiness of this practice, the parties to a transaction, or their successors, almost never requested that a complete record be issued: the existence of the *imbreviatura* in a notary archives was sufficient evidence of the transaction.

This practice is illustrated in Figure 1 below. The Formulary includes invariant content data, articulates the essential aspects of documentary form, and indicates the variable data elements whose values must be specified in each instance, while the Imbreviatura contains the instance data for a single transaction. Thus, the imbreviaturae for all transactions of type "a" would include the data elements specified in the formulary for transactions of that type. In the imbreviaturae system, records of various transactions, such as contracts and deeds, were kept, but not in the documentary forms which were prescribed for those transactions. The record keeping system enabled the record keepers to produce copies of the records in the required forms on demand. Any instance of a transactional record produced in this manner was an original.[59] Interestingly, the imbreviatura system was so reliable that, over time, people who needed to know what was in a transactional record were satisfied by ascertaining the data contained in the imbreviatura, and did not require the production of records in the documentary forms prescribed in the applicable formulary. Even though the imbreviatura was not the intended complete record, but only a prelude to it, it served as a reliable record of a transaction of a given type. The reliability was contingent on the fact that the system could produce an authentic complete record of the transaction on demand.

Functionally, the formularies and the imbreviaturae are the equivalents of digitally stored records described above.[60] While the imbreviaturae system physically separated content data about a specific transaction from the documentary form and the invariant content

---

[59] Technically, the very first was an original, while the subsequent ones, lacking primitiveness, were copies in the form of original, although they had the force of an original. However, as they were all produced directly from the imbreviatura rather than from one another, and, having a different date of transmission and, possibly, a different writer (a notary who has legitimately succeeded to the original one) and a different addressee (the addressee of the action would remain the same, but the addressee of the record could be a descendant of one of the original parties), were different records, they can be all regarded as originals.

[60] There is a subtle, yet important, distinction here that makes the correspondence to the Medieval analogy less than exact. The digital records are not being stored in documentary forms that are different than originally *intended*, rather, they are being stored in forms that are different than the forms in which they were originally *created*. The imbreviaturae (notwithstanding that they are complete records unto themselves) are, in one sense, incomplete 'stand-in' records for complete transaction records that, in most cases, were never created. This is not the same process that is occurring with the digital records, because the digital records are, in fact, created as complete records in their final form, and then saved in a form that differs from the form in which they were originally created. Thus, unlike the potential 'records to be' that are associated with the imbreviaturae system, the products of the digital systems discussed here are complete records created in their final form prior to being stored (except in those cases, as noted, where the minimum requirements for an e-record are not met).

(i.e., extrinsic and intrinsic elements of form), physical separation is but one of many possibilities in the digital environment. Digital information technology offers a variety of ways of keeping and combining data and instructions. What is essential is that the computer stores and processes the data and the instructions in a way that consistently and correctly distinguishes each type and combines the different digital components of a record.[61]

Similar practices exist in the digital environment. Figure 2, depicts in the abstract a common way of keeping records using database applications. The three elements at the top of the figure, "Database Form," "Instance Data" and "Document Copy," illustrate in summary fashion how individual documents can be generated by applying a Database Form to the data of individual instances of the types of transactions covered by the form. This process exactly parallels the medieval use of imbreviaturae. However, digital information technology offers greater variations in how this practice can be implemented. The lower part of Figure 2 presents a more detailed view of how forms are filled in with data of individual instances. In a database, not all of the data need to be kept in a single 'document' or logical object within the application. A form can be filled in with data of a single instance stored in many different locations in the database. Figure 2 illustrates this for a relational database, where the data of each individual case are spread across several tables. The database keeps track of the data of individual instances thanks to a logical data model that defines how data entries in one table are related to those in other tables. The database application fills in a form, ensuring that the right data elements of a single instance are entered in the correct locations on the form in accordance with rules that map the logical data model to the form. An additional element of complexity in the digital environment is that individual data can be combined in different ways with other data to produce different types of documents; for example, data from an online order can be used to fill out a pull list, shipping label, and invoice.

In this conceptualization of electronic records, when is a digital component of a manifested record a record itself? One example of such a situation occurs when a single digital component, such as a binary image of a printable document, comprises all of the data necessary to reproduce the manifested record: in this case there is a

[61] Future analysis of how the different types of data are mapped into digital components should address how the system recognizes and processes the different data types in different mappings.

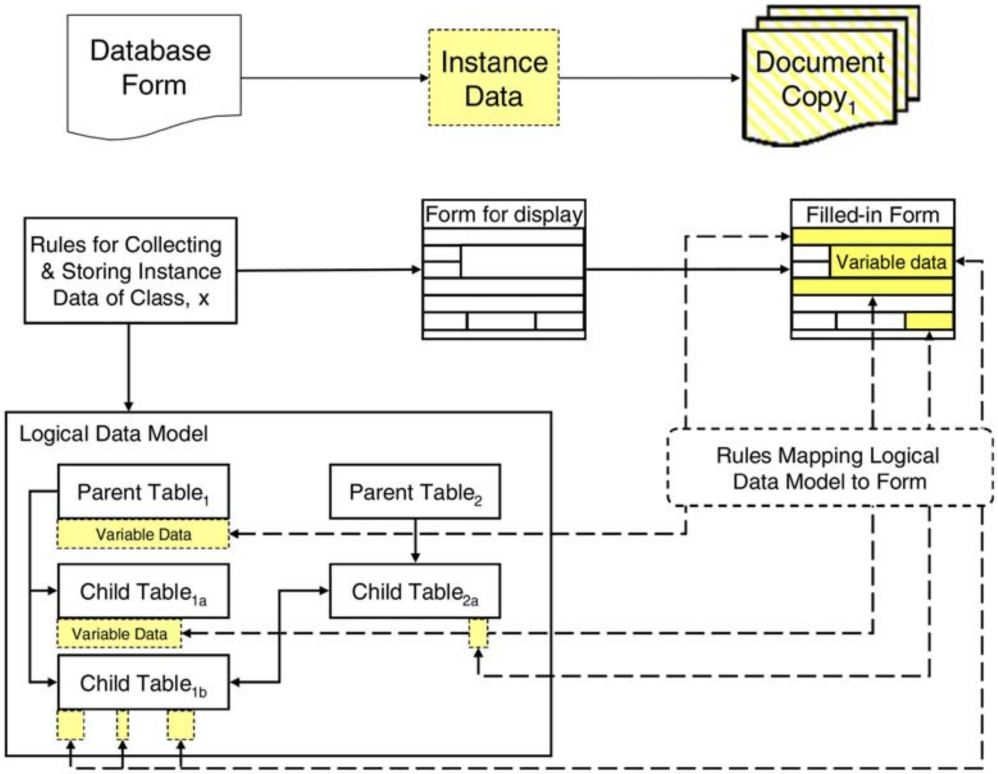


Figure 2. Producing Documents From Databases.

one-to-one correspondence between the digital component and the manifested record. Another example occurs when, in a complex database like that illustrated in Figure 2, one digital component, the "database form," has a fundamental unity, autonomy and completeness, just like a medieval formulary. This would not be the case for the digital component "instance data," because each of the data necessary to fill in the form is recognized and processed by the database application as a distinct bit stream, and is therefore meaningless by itself. It would acquire meaning only in the context of a data model. Moreover, while the application can isolate each datum, it does not store or manage it as a distinct object. Data are defined and stored as part of database tables. Thus, if properly managed, each database table, the logical model of the database, and any other model, such as one that defines a form or report, may constitute a stored record.

Figure 3 presents a generalized model summarizing the analysis of electronic records to this point. It can be applied to traditional, interactive, experiential and dynamic computing environments that produce some type of manifested document, and it can be used as an

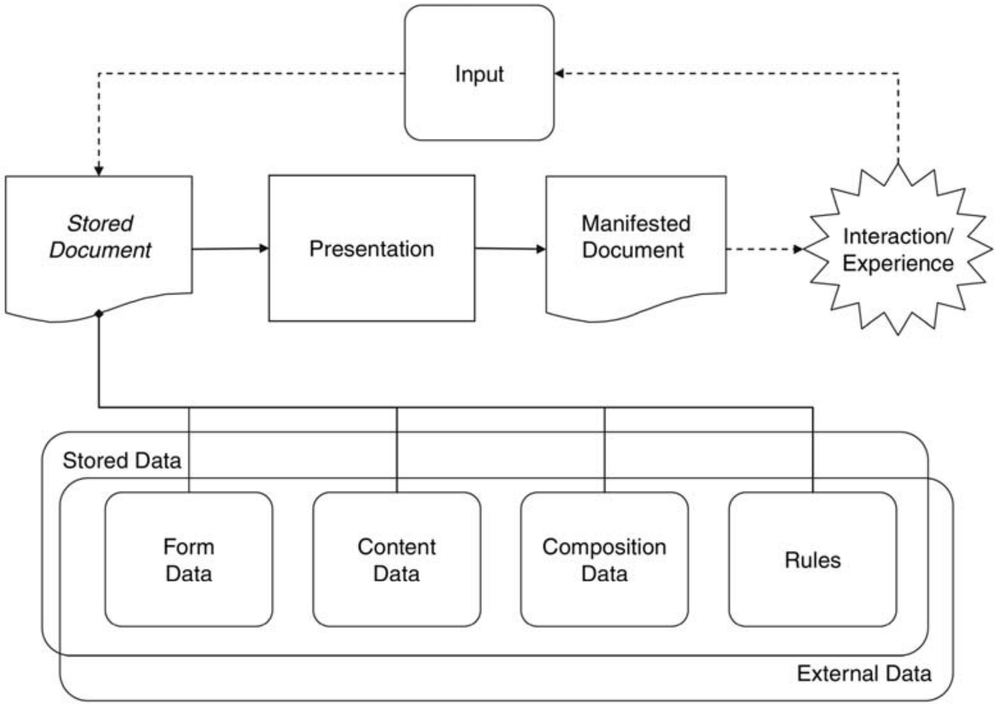


Figure 3. Generic Model of Stored and Manifested Documents.

analytic tool to determine the characteristics and nature of electronic records.[62] The first three figures, reading from the left, in the middle row of Figure 3, depict in very abstract fashion the production of a Manifested Document from a Stored Document, a process common to all these environments. In a system that does not have interactive or experiential features, the final figure in the middle row, Interaction/Experience, would not occur, hence the dotted lines connecting this figure. The Manifested Document may be presented to a person or another system. It is assumed to include all the content that could be manifested, even if only a portion of the content is manifested at one time, as in the case of audio and video recordings and pseudocs. In order for the Manifested Document to be a record it must be possible to reproduce it repeatedly as it appeared the first time. If the environment provides for Input (top row) from a user or interacting system and that input can change the content or form of the Manifested Document, then this document cannot be a record even if other

[62] Such analysis, however, would only be partial. It is also necessary to determine, through additional analysis, whether the requirements related to context, action, persons, archival bond, and intrinsic elements of form are satisfied.

requirements for being a record are satisfied. Regardless of its record nature, though, there may be one or more digitally stored records used to produce the Manifested Document. If the system stores an object that is the internal representation of the Manifested Document, that object may be a stored record. In the case of a static Manifested Document, the Stored Document should contain all of the Content and the Form Data and Rules that determine the extrinsic elements of the form of the Manifested Document. Where there is a one-to-one correspondence between the Stored and Manifested Document, there would be no Composition Data because the Stored Document is already composed. The Stored Document would include provisions for modifying the content and/or form of the Manifested Document, but the Stored Document might still be a record, analogous to the medieval formulary, if it presents all the other necessary characteristics of a record. However, if the Stored Document is itself modifiable as a result of Input or External Data, it cannot be a record. If the system does not store a single representation of the Manifested Document, it is necessary to determine how it composes that document from Form Data, Content Data, Composition Data, and related Rules and whether the digital components which make up those types of data are stored entirely within the system or derived in whole or in part from one or more external sources, in order to identify any stored records.

A record is whatever the creator treats as its record, but that "whatever" must be something that the creator can in fact keep, associate with other records, and subsequently recall. There are two different modes in which a record can serve a memorial function. In most cases, the memorial function of a record is retroactive: it is the means through which its creator remembers what was done, happened, or was described or said, and through which others may learn about the past. For example, the record of a performance of an artistic work is retrospective: it enables the audience to remember – or more accurately in most cases to experience – how the artwork was executed in that specific performance. But, there are also records whose principal function is prospective. A musical score or the script for a play may be considered a record of the artist's career or genius. It can be examined for what it reveals of the evolution of the artist's abilities and leanings, of the impact of biographical events on artistic output, etc. However, the driving intent in the creation of the play or piece of music is that it be performed. The script or score serves a prospective function: it is a set of instructions on actions to be carried out afterwards.

The distinction between retrospective and prospective records can help us come to grips with records in interactive, experiential and dynamic environments. *Interactions* between humans and computer systems, *experiences* enabled or mediated by experiential systems, and *processes* which are composed and carried out with at least some degree of spontaneity by dynamic systems are not the residues of action. They are not means of remembering either what was done or what is to be done. In short, they are not records. But, they can be captured in documentary form and some of these documents could be treated and used as records of interactions, experiences, or dynamic processes, that is, they may become records of those activities. In addition, interactions, experiences and processes are enabled by documents within such systems and these documents can serve as prospective records. Both retrospective and prospective documents can be found in all three-focus areas examined by InterPARES 2. In the arts, there are recordings of performances and documents that enable performances. In government, documents created in the execution of governmental transactions can be retrospective records. In science, documentation of the conduct and results of experiments and observations are retrospective. In government, laws, regulations, and directives and, in science, research plans and protocols are created with the primary intention of guiding, controlling, or perhaps prohibiting subsequent actions.[63] In sum, retrospective records capture, while prospective records enable or at least inform interactions, experiences, or dynamic processes.

Within the class of prospective records, there are two subclasses. One simply contains instructions about executing an action or process. The other subclass is actively involved in carrying out the action or process. Examples of "instructive" records include musical scores, regulations, manuals of procedures, and instructions for filling out forms. Examples of "enabling" records include software patches that enable a musical instrument to interact with a computer, software in online marketing sites that interprets data about a visitor's actions on the site to determine what elements of content should be presented next to that visitor, and software agents that enable interacting business applications to execute transactions autonomously. Although software is not commonly considered a record, rather it is to be

[63] Prospective records still retain the basic function of remembrance: they enable subsequent actions and actors to remember what to do and/or how to do it in accordance with prior decisions.

regarded as a digital component of records, this type of software is created and serves as a record in the specific contexts presented here, as it is generated and used as a means for carrying out the specific activity in which it participates and stands as the instrument, byproduct and residue of that one activity.

In addition to the differences in how they are involved in the actions or processes that they inform or control, there are differences in the way instructive and enabling records are materialized to achieve the purpose for which they are created. Instructive records are intended to be read by humans and, therefore, are materialized by being reproduced from stored digital components into a human readable form. In contrast, enabling records achieve their purpose in the digital form in which they are stored[64] and, conversely, cannot achieve that effect if transformed into human readable format. Moreover, as long as they remain active, enabling records must be maintained in the systems in which they were created – or in systems with identical functionality. Otherwise, they will not produce or enable the interactions, experiences, performances or other processes they were intended to generate.[65]

Keeping Interactive, Experiential, and Dynamic Records

Digitally stored records are kept in order to be able to reproduce manifested records. There are three broad possibilities for keeping stored records: (i) keeping instructive records in the system in which they are generated, together with all the instantiations generated from them, (ii) keeping instructive records in another system, and (iii) keeping enabling records and the record instantiations produced from

[64] Strictly speaking, computer code is not stored in the form in which it controls or shapes processes. It needs to be translated into machine code at the time of execution, but that translation is analogous to the translation of a musical score into signals processed by the human brain during performance of the work.

[65] The situation is reversed when such records become inactive. In order for a human to understand what these records did in their original technological, documentary and administrative contexts, it is necessary to convert them from the form in which they were stored and functioned as records to a form that humans can read; for example, instructions must be converted from the binary form in which they were executable to a textual form. In most cases, this conversion will involve translation from machine language to a humanly readable one.

them in the system in which they are used.[66] Of course, the possibility of keeping records does not entail that records actually are kept. Individual cases need to be examined to determine if and how records are kept.

The first possibility for keeping digitally stored records is to retain them in the interactive, experiential, or dynamic systems in which they are generated. Given that a system has the ability to produce the manifested records in the first place, it could be designed to reproduce them subsequently from stored records. An example of an instructive record kept in an interactive system is that of the script for Waking Dream, which is kept on the website for the work. The Alsace-Moselle Land Registry is an example of a system which is used both to carry out transactions and to keep the records of those transactions.

The possibility of reproducing a record using the same functionality that produced it in the first plase does not exist in cases where some of the content is not stored in the system. In the case of record instantiations, even if the system has the capability to produce the same document in response to the same input over and over again in a reliable way, this does not necessarily mean that the system keeps the document instantiation as a record. This difficulty can be seen in an InterPARES 2 case study in the scientific focus, the CyberCartographic Atlas of Antarctica. D.R.F. Taylor defines cybercartography as "the organization, presentation, analysis and communication of spatially referenced information on a wide variety of topics of interest and use to society in an interactive, dynamic, multimedia, multimodal and multidisciplinary format." The CyberCartographic Atlas of Antarctica incorporates scientific and environmental data into "an online atlas portraying, exploring and communicating the complexities of the Antarctic continent for education, research and policy purposes."[67]

---

[66] The InterPARES 1 report suggested another possibility, that of trading the record characteristics of stability of content and fixity of form (including completeness of content and form with respect to the first and to any subsequent instantiations of the record) with the ability of the system containing it to track and preserve any change to the record. See http://www.interpares.org/book/interpares_book_d_part1.pdf, p.24. In other words, the researchers were inclined to shift the requirements of stability and fixity from the record to the log of the changes to the record once the record was no longer active; in this context, the object identified as the record and to be kept intact would then be the last instantiation of the fluid object, plus the complete log of changes, and the metadata of both. This option is conceptually sound only if the creator uses this set of objects as its record, but this scenario is very unlikely because it would be highly impractical.

[67] Tracey P. Lauriault, Peter Pulsifer and D.R. Fraser Taylor. The Cybercartographic Atlas of Antarctica Project. See http://www.carleton.ca/gcrc/caap.

The Atlas organizes heterogeneous data on physical, biological and human influenced characteristics and their interactions into content modules along thematic lines, and supports a variety of discovery, visualization and access functions. Projected data volumes are large and expected to grow over time. Data values in the Atlas are in stable and basic formats defined to support different presentations appropriate to experts, the general public and policy makers. However, these formats are essentially starting points enabling users to explore and access the contents interactively. The Atlas user can select among a rich and diverse, but still finite set of tools for searching, visualizing, hearing or otherwise accessing heterogeneous data about topics of interest. Hence, the forms in which the data are presented are not immutable but, as with works of performance art, their variability is within parameters established by the author. In principle, every user of the Atlas making the same selection of content and form of presentation would see an identical document; therefore, the Atlas virtually keeps such documents. However, they do not qualify as record instantiations because the system does not retain any data about their production. There is no chronological date, no identification of the addressee, nor any information about the activity in which the document is first produced.

A stored dynamic document, such as one which stores user inputs and uses them in subsequent manifestations or one which processes and presents, but does not store, data from users or from other external sources, might be said to be always in the process of creating, but never completing a manifested record. The manifested document might be a record if the processes which cause it to be forever in progress were terminated, or if it were removed from the dynamic environment and kept in some frozen form. However, the final state of a document somehow isolated from or rendered immune to dynamic processes would be static. But producing a static document in either way would amount to creating a different document. It might serve as a record *of* the dynamic process or its state at the moment it was frozen, but it would not be able to serve the dynamic purpose of producing variable output in response to a variety of different inputs or stimuli. In some cases, the stored dynamic document might be kept as a record, but that would not be possible in the case of documents covered by class 2.2.4 of Table I.

The difficulty of keeping dynamic documents as records is found in the VanMap case study. VanMap does not keep records, but could be modified to do so. As long as data are overwritten by updates, their aggregation as it appears at any given time will never reach the state of a document, let alone a record. The combination of data layers, sets of data, and set of instructions which produce displays used in any business process can be described as an enabling document of that process. The system may preserve these documents for some time: it could reproduce the information used in a particular action, in the same form and with the same content as when the action was carried out, but only as long as there is no modification of either the data or the instructions which control the form of display. It should be possible to modify the design and operation of VanMap either to output records or store a stratification of the data variations over time.

Being able to keep records in the system used to carry out transactions is a necessity in cases where the transactions involve multiple successive steps and where the interacting systems support multiple transactions in which records of earlier transactions are needed in subsequent ones. An interacting system could not proceed to the next step or the next transaction if it did not contain records of the previous step or transaction.[68] That is not to say that the system must be able to reproduce the records using the same functionality that it used to produce them in the first place. The system could be designed to include a record keeping subsystem in addition to the subsystem or module used to produce the records.

The second possibility for keeping records generated in interactive, experiential, or dynamic systems is to retain them in a system specifically designed for record keeping. This could be achieved either by taking them out of the original system and retaining them in another system suitable for record keeping, or by adding record keeping functionality to the original system. Given that record instantiations and instructive records are outputs of such systems, they could be kept in some appropriate digital format as either static records or records with bounded variability. As that the InterPARES 2 project has examined interactive, experiential and dynamic systems, but not record keeping systems, there are no case studies of separate systems for keeping electronic records. The VanMap system does output record instantiations to other systems, but they are hard copy outputs.

Keeping electronic records in a separate system or subsystem designed specifically for this purpose is necessary in any case where the

[68] It might be argued that such a system does not necessarily need to maintain records of earlier steps in a process or earlier transactions, but that it would suffice if the system kept the data about those steps or transactions. However, if such a system does not demonstrably satisfy requirements for keeping records, it should not be relied upon to carry out business over time.

system which produces the records does not have the capability to reproduce the manifested records reliably using the original functionality which produced them. But a document set aside and kept as a record in this way may have some differences in both content and form compared to the original output. This record would not be identical to the actual system product, precisely because it has lost whatever interactive, experiential or dynamic properties the product had. However, one should consider that all records are substitutes: they stand for or take the place of acts or facts. The ability to reproduce what the system presented in a particular instance, without the interactive, experiential, or dynamic attributes of the environment in which it was produced, may be sufficient for some record keeping needs.

One method of keeping records in this way would be to take a snapshot or otherwise freeze a presentation by a dynamic, experiential, or interactive system. This is done, for example, with a static audio recording of a performance of music or audio/visual recording of other performance art. In such cases, the static record would represent, but not identically repeat the original performance.

Static, durable objects may in fact be necessary to satisfy the creator's needs for records. An example of an interactive system which creates static records as it interacts with users to meet the creator's record keeping needs is that of an online sales application which accumulates the data it needs to produce purchase orders from user selections and inputs. Another example might be a dynamic system used to execute a manufacturing process. It may record the variations – including both different values of data and different processes executed – which occur in each instance of execution, and output that information in a report. In these examples, the static records are records of transactions the system is designed to support. Given that the CyberCartographic Atlas of Antarctica is designed to provide information rather than to support transaction of business, any document output from the Atlas cannot be a record with respect to the system, but it could subsequently be used by the recipient in its own business. A user could create a record by collecting and acquiring content elements which could be exported from the Atlas to the user's environment or, if applicable, printed.

A single system may employ both possibilities for keeping records described thus far, regardless of whether it supports interactions with other systems or with human users. For example, while a system executing transactions through interactions with other systems will retain copies of the sent documents as records, and set aside received documents as records, thereby ensuring that all records are in the system in which they are created, the Alsace-Moselle Land Registry makes use of both the first two possibilities for keeping records. It reproduces records of active entries using the same XML and database capabilities used to create new real property entries, and it has been extended to keep and retrieve the scanned image records of inactive entries.

The third possibility for keeping electronic records applies to enabling records. Given that such records are instrumental in achieving the outcomes or producing the outputs intended by a system or application, it is likely that an enabling record will be maintained in the system for as long as the system itself is maintained.

As a whole, the CyberCartographic Atlas can be described as an interactive and experiential enabling record of the Geomatics and Cartographic Research Centre at Carleton University, made in the course of a research project funded by the Social Sciences and Humanities Research Council of Canada, that shares an archival bond with the project's administrative records.

However, interactive, experiential, or dynamic systems do not necessarily contain enabling records. In dynamic systems, even in cases where it is possible to identify a set of digital components which might appear to be an enabling record, if the system has the capability of adapting its software autonomously, the object comprising those digital components will not satisfy the basic requirements of fixed content and form. In such cases, there is no way to recall what was the process which produced a given outcome or output. Thus, the system in question needs to be modified.

In VanMap, objects that exist in the system might be maintained to provide records reflecting the situation observed by the decision makers at any given time. Following well established database management methods, a history file of the data in the system could be created and set aside before any update, along with a detailed description of each business process in which VanMap is involved and of the way in which VanMap is used in each of them, in addition to the instructions used to create the records supporting each type of transaction. The description would reveal the archival bond between the records of each business process and VanMap and the instructions would reveal the specific relationship between each process and the data which supported it. This approach would follow a centuries-long tradition of embedding in a code of administrative procedure the function of a record that serves multiple activities and procedures, but of which only one original exists (see for example the series of the maps of the cadastre, which were and are used as records in several procedures having different purposes). However, as currently designed and operated, VanMap does not preserve records of any business transactions.

A fixed form and a stable content are only two of the characteristics that an electronic entity must present in order to be considered a record. Entities that are records must also have explicit linkages to other records, an identifiable administrative context; an author, an addressee, and a writer; and an action, in which the record participates or which the record supports either procedurally or as part of the decision making process. But, this an area of analysis and discussion for another article.

Conclusion

The InterPARES project has examined a panoply of topics concerning the preservation of authentic records in electronic systems. Among these topics are the characteristics of electronic records. Applying and testing traditional concepts drawn from diplomatics and archival science to a considerable number of case studies, the project has studied how records in electronic systems – when they exist – resemble and differ from traditional records in hard copy. In the first phase of the project, the case studies focused on the digital counterpart of traditional records. The most salient empirical characteristics of such records are that their digital encoding does not manifest the documentary form of the record and that, therefore, they are not preserved as physical objects, but as one or more bit streams that must be correctly processed by computers to be rendered in the proper documentary form. In its second phase, InterPARES is examining interactive, experiential and dynamic systems that do not necessarily produce or keep anything that corresponds to traditional records.

The analysis of interactive, experiential and dynamic systems requires clear distinctions among the systems themselves, the interactions, experiences, performances and other outputs, and the objects generated and/or kept in them. In many cases, interactive, experiential and dynamic systems produce objects that have the appearance of documents, but, after their first manifestation, cannot be re-produced with the same content and in the same form. Given the essential memorial function of a record – a record is a residue of activity retained by its creator for reference or use in later activity – such cases appear to fail the basic requirement of a fixed form and stable content.

However, cases discussed in this article show that interactive, experiential and dynamic systems can produce documents capable of being kept as records. Furthermore, a closer examination of these systems shows that they may contain documents which exhibit some variability in form and content, but, because their variability is appropriately bounded, can be considered records, as when the variability is due to technology rather than to the intention of the author or writer of the document. In addition, authors or writers can generate digital records that embed intentional variability. This includes the construction of documentary forms that enable users to select subsets of content and control both the sequencing and presentation features, such as image magnification.

There are also cases, most notably in the arts, but also in government and science, where interactive, experiential and dynamic systems contain documents whose presentation or rendering always shows some unique or spontaneous variation in content or form. In such cases, one must distinguish between what is output by the system and the document(s) that enable the system to produce its output. Such documents are 'enabling:' they enable the interactions, experiences or processes the system executes. Provided they are properly maintained and managed as intellectually interrelated parts of records aggregations, enabling documents can be considered records. On first encounter, this conclusion seems to contradict the finding of InterPARES 1 that an electronic record is not something kept in a system, but something reproduced by processing data objects stored in the system, but, as demonstrated, this conclusion broadens, rather than contradicts, the finding of InterPARES 1.

Of course, defining the concept of record in the context of interactive, experiential and dynamic systems is a very tall order, primarily because it depends on the perspective (i.e., whether one looks at the digital entities from the point of view of the author/writer, the user, the creator, or the preserver) and on the particular level of abstraction in question (i.e., whether at the entire-object-as-record level, on down to the individual-object-interactions-as-record level). However, the purpose of this article was to present the work done by InterPARES in this respect and to initiate a theoretical discussion. In spite of its length, this article is only a beginning of that discussion. The concept of a record here articulated needs to be tested in other environments. The practical possibilities of preserving such records needs to be explored, notably in the context of the major contributions to digital preservation outside of the InterPARES project, such as the OAIS model, the CEDARS and the CAMiLEON[69] projects, the METS standard,[70] and PREMIS metadata. There is an undoubted

[69] See http://www.si.umich.edu/CAMILEON/.

[70] See: http://www.loc.gov/standards/mets/.

need to explore the great practical implications and legal consequences for all the parties directly involved, and all the stakeholders. It is the hope of these authors that such discussion will continue well beyond the conclusion of the InterPARES project.

Appendix

Appendix: Glossary of Key Terms

Term	Definition
Composition data	One of the three types of stored digital data used to produce or reproduce a digital document, they tell the system what form and content data belong to which document.
Content data	One of the three types of stored digital data used to produce or reproduce a digital document, they constitute the content of a document.
Digital component	A bit stream that is necessary to reproduce the document. It may comprise composition, content or form data, or some combination of such data.
Form data	One of the three types of stored digital data used to produce or reproduce a digital document, they enable the system to reproduce the document in correct form.
Manifested record	The visualization or materialization of the record that is produced from the stored digital component(s) in a form suitable for presentation to a person or another system.
Presentation features	A set of perceivable features (graphic, aural, visual) generated by means of encoding and program instructions, and capable, when used individually or in combination, to present a message to our senses.
Stored record	A digitally encoded object which is managed as a record.

APPENDIX 3

InterPARES 2 Case Studies and General Studies: Researchers, Focus and Final Status

Case Studies

CS01 Arbo Cyber, théâtre (?)

Focus:	1 (Arts)
Status:	Completed
Principal investigator:	Martine Cardin (archivist), Université Laval[1]
Graduate research assistants:	Philippe Perron, Université Laval Carolyn Petrie, The University of British Columbia

CS02 Performance Artist Stelarc

Focus:	1 (Arts)
Status:	Completed
Principal investigator:	Henry Daniel (dancer), Simon Fraser University
Study team member:	Luciana Duranti (archivist), The University of British Columbia
Graduate research assistants:	Peggy Heger, The University of British Columbia Cara Payne, The University of British Columbia

CS03 *Horizon Zero*/Zero Horizon Online Magazine and Media Database

Focus:	1 (Arts)
Status:	Completed
Principal investigator:	Brent Lee (musician), University of Windsor
Study team member:	Justine Bizzocchi (technologist), Banff Centre

CS04 Persistent Archives Based on Data Grids

Please see General Study 01.

CS05 Archives of Ontario Web Exhibits

Focus:	3 (Government)
Status:	Completed
Principal investigator:	Jim Suderman (archivist), Archives of Ontario
Study team member:	Marta Braun (art historian), Ryerson University Barbara Craig (archivist), University of Toronto Michael Murphy (technologist), Ryerson University
Graduate research assistants:	Deidre Brocklehurst, The University of British Columbia Terra Dickson, The University of British Columbia Peggy Heger, The University of British Columbia Brenda McPhail, University of Toronto

CS06 Cybercartographic Atlas of Antarctica

Focus:	2 (Science)
Status:	Completed
Principal investigator:	Fraser Taylor (geographer), Carleton University
Study team member:	Yvette Hackett (archivist), Library and Archives Canada
Graduate research assistant:	Tracey Lauriault (case study lead), Carleton University

[1] Institutional affiliations changed over the course of the Project. This appendix generally lists the researcher affiliations at the time the specific study was conducted.

CS08 Mars Global Surveyor Data Records in the Planetary Data System

Focus: 2 (Science)
Status: Completed
Principal investigator: William Underwood (technologist), Georgia Tech Research Institute

CS09 Digital Moving Images: Inputs, Processes and Outputs (a case study in four parts)

Focus: 1 (Arts)
Status: Completed
Principal investigator: James Turner (information specialist), Université de Montréal
Study team members: Marta Braun (art historian), Ryerson University
Mary Ide (archivist), WGBH Boston
Randal Luckow (archivist), Turner Broadcasting System
Michael Murphy (technologist), Ryerson University
Andrew Rodger (archivist), Library and Archives Canada

CS09(1) Altair4 di Roma. A Multimedia Archaeological Project: The House of Julius Polybius

Principal investigator: Isabella Orefice (archivist), Associazione Nazionale Archivistica Italiana

CS09(2) National Film Board of Canada

Principal investigator: Andrew Rodger (archivist), Library and Archives Canada

CS09(3) Commercial Film Studio

Co-Principal investigators: Randal Luckow (archivist), Turner Broadcasting System
James Turner (information specialist), Université de Montréal

CS09(4) WGBH Boston

Principal investigator: Mary Ide (archivist), WGBH Boston

CS10 *The Danube Exodus*: Interactive Multimedia Piece

Focus: 1 (Arts)
Status: Completed
Principal investigator: Sally Hubbard (archivist), Getty Institute

CS11 Nova Scotia Business Registry, Service Nova Scotia and Municipal Relations

Focus: 3 (Government)
Status: Retired
Principal investigator: Margaret Campbell (archivist), Archives of Nova Scotia

CS12 Antarctic Treaty Searchable Database

Focus: 3 (Government)
Status: Completed
Principal investigator: Paul Berkman (environmental scientist), Ohio State University/EvREsearch, Ltd.
Study team members: Babak Hamidzadeh (technologist), Library of Congress
Richard Marciano (technologist), San Diego Supercomputer Center

		Reagan Moore (technologist), San Diego Supercomputer Center George Morgan (technologist), EvREsearch, Ltd. Jim Suderman (archivist), Archives of Ontario
CS13	***Obsessed Again…***	
	Focus:	1 (Arts)
	Status:	Completed
	Co-principal investigators:	Keith Hamel (musician), The University of British Columbia Jesse Read (musician), The University of British Columbia
	Study team member:	Luciana Duranti (archivist), The University of British Columbia
	Graduate research assistant:	J. Scott Amort, The University of British Columbia
CS14	**Archaeological Records in a GIS: Research in the American Southwest**	
	Focus:	2 (Science)
	Status:	Completed
	Principal investigator:	Richard Pearce-Moses (archivist), Arizona State Library
	Graduate research assistants:	Erin O'Meara (case study lead), The University of British Columbia Randy Preston, The University of British Columbia
CS15	***Waking Dream***	
	Focus:	1 (Arts)
	Status:	Completed
	Principal investigator:	Sidney Fels (technologist), The University of British Columbia
	Study team member:	Luciana Duranti (archivist), The University of British Columbia
	Graduate research assistant:	Seth Dalby, The University of British Columbia
CS16	**Model for Description and Preservation of Documents Created Unstable and Variable Artistic Techniques**	
	Focus:	3 (Government)
	Status:	Completed
CS17	**New York State Department of Motor Vehicles On-line Services System**	
	Focus:	3 (Government)
	Status:	Completed
	Co-principal investigators:	Phil Eppard (archivist), State University of New York, Albany Terry Maxwell (archivist), State University of New York, Albany Mark Wolfe (archivist), State University of New York, Albany
	Graduate research assistants:	Joshua Hauck-Whealton, State University of New York, Albany Richard Hoppenstedt, State University of New York, Albany

		Rachel McMullin, State University of New York, Albany Peter Runge, State University of New York, Albany Reginald White, State University of New York, Albany Jessica Zacher, State University of New York, Albany

CS18 Computerization of Alsace-Moselle's Land Registry

Focus:	3 (Government)
Status:	Completed
Principal investigator:	Jean-François Blanchette, The University of British Columbia
Graduate research assistant:	Geneviève Shepherd (case study lead), The University of British Columbia

CS19 Preservation and Authentication of Electronic Engineering and Manufacturing Records

Focus:	2 (Science)
Status:	Completed
Principal investigator:	Kenneth Hawkins (archivist), National Archives and Records Administration

CS20 Revenue On-Line Service (ROS)

Focus:	3 (Government)
Status:	Completed
Principal investigator:	John McDonough (archivist), National Archives of Ireland
Study team members:	Ken Hannigan (archivist), National Archives of Ireland Tom Quinlan (archivist), National Archives of Ireland

CS21 Electronic Filing System (EFS) of the Supreme Court of Singapore

Focus:	3 (Government)
Status:	Completed
Principal investigator:	Elaine Goh (archivist), National Archives of Singapore

CS22 Electronic Café International (ECI)

Focus:	1 (Arts)
Status:	Retired (some reports available)
Principal investigator:	Howard Besser (moving images specialist), New York University
Graduate research assistants:	Nadine Hafner, The University of British Columbia Janine Johnston, The University of British Columbia Tracey Krause, The University of British Columbia Keum Hee Yu, The University of British Columbia

CS23 UK Knowledge Network

Focus:	3 (Government)
Status:	Retired (some reports available)

CS24 City of Vancouver Geographic Information System (VanMap)

Focus:	3 (Government)
Status:	Completed
Principal investigator:	Evelyn Peters McLellan (archivist), City of Vancouver
Study team members:	Sue Bigelow (conservator), City of Vancouver Archives

	Glenn Dingwall (archivist), City of Vancouver Archives Luciana Duranti (archivist), The University of British Columbia Richard Marciano (technologist), San Diego Supercomputer Center Jonathan Mark (technologist), City of Vancouver, Information Technology Department Andrew Power (records manager), City of Vancouver Peter Van Garderen (archivist), President, Artefactual Systems, Inc. Reuben Ware (archivist), City of Vancouver Archives Liz Wright (archivist), City of Vancouver Archives
Graduate research assistants:	Eleanor Kleiber, The University of British Columbia Catherine Miller, The University of British Columbia

CS25 Legacoop of Bologna Web Site

Focus:	3 (Government)
Status:	Completed
Principal investigator:	

Isabella Orefice (archivist), Associazione Nazionale Archivistica Italiana

CS26 Microvariability & Oscillations of Stars (MOST) Satellite Mission: Preservation of Space Telescope Data

Focus:	2 (Science)
Status:	Completed
Principal investigator:	Reagan Moore (technologist), San Diego Supercomputer Center
Study team member:	Luciana Duranti (archivist), The University of British Columbia
Graduate research assistant:	Bart Ballaux (case study lead), The University of British Columbia

General Studies

GS01 Persistent Archives Based on Data Grids

Focus/Area:	Preservation
Status:	Completed
Principal investigator:	Reagan W. Moore (technologist), San Diego Supercomputer Centre

GS02 Survey and Analysis of Scientific Encoding Languages for Non-Textual Records

Focus/Area:	Preservation
Status:	Retired (some reports available)
Principal investigator:	William Underwood (technologist), Georgia Tech Research Institute

GS03 Preserving Interactive Digital Music - the MUSTICA Research Initiative

Focus/Area:	1 (Arts)
Status:	Completed

Co-principal investigators:	Jean-François Blanchette, The University of British Columbia John Roeder (musician), The University of British Columbia
Graduate research assistants:	Jill Teasley (case study lead), The University of British Columbia Jennifer Douglas, The University of British Columbia Carolyn Petrie, The University of British Columbia

GS04 Survey of Recordkeeping Practices of Composers

Focus/Area:	1 (Arts)
Status:	Completed
Principal investigator:	Michael Longton (musician), University of Victoria
Graduate research assistant:	Vincent Schillaci-Ventura, The University of British Columbia

GS05 An Examination of the Processes to Preserve and Manage Electronic Records: Round Three at The National Archives of Australia and WGBH

Focus/Area:	Preservation
Status:	Retired (some reports available)
Co-principal investigators:	Shelby Sanett (archivist), U.S. National Archives and Records Administration Michèle V. Cloonan (information specialist), Simmons College

GS06 A Bayesian Belief Network: Supporting the Assessment of the Degree of Belief that a Recordkeeping System Maintains Authentic Digital Records

Focus/Area:	2 (Science)/Appraisal
Status:	Completed
Co-principal investigators:	William Underwood (technologist), Georgia Tech Research Institute Sheila Isbell (technologist), Georgia Tech Research Institute

GS07 Recordkeeping Practices of Photographers using Digital Technology

Focus/Area:	1 (Arts), 2 (Science), 3 (Government)
Status:	Completed
Principal investigator:	Marta Braun (art historian), Ryerson University
Study team member:	Yvette Hackett (archivist), Library and Archives Canada
Graduate research assistant:	Jessica Bushey, The University of British Columbia

GS08 Survey of Government Web Site Interactivity

Focus/Area:	3 (Government)
Status:	Completed
Principal investigator:	Jim Suderman (archivist), Archives of Ontario
Study team members:	Yvette Hackett (archivist), Library and Archives Canada Sue McKemmish (archivist), Monash University Mark Wolfe (archivist), State University of New York, Albany

GS09 Digital Recordkeeping Practices of GIS Archaeologists Worldwide

Focus/Area:	2 (Science)
Status:	Completed
Principal investigator:	Randy Preston (archivist), The University of British Columbia
Study team member:	Erin O'Meara (archivist), The University of British Columbia

GS10 Preservation Practices of Scientific Data Portals

Focus/Area:	2 (Science)
Status:	Completed
Principal investigator:	Barbara Craig (archivist), University of Toronto
Graduate research assistants:	Heather Dean, The University of British Columbia Stephen Gage, The University of British Columbia Erin Hanlon, The University of British Columbia Tracey Lauriault, Carleton University Christina Miller, The University of British Columbia Brian K. Trembath, The University of British Columbia Sherry Xie, The University of British Columbia

GS11 Selecting Digital File Formats for Long-Term Preservation

Focus/Area:	Preservation
Status:	Completed
Principal investigator:	Evelyn Peters McLellan (archivist), Insurance Corporation of British Columbia
Graduate research assistants:	Tracey Krause, The University of British Columbia Yvonne Loiselle, The University of British Columbia

GS12 Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data

Focus/Area:	Creation, maintenance, preservation
Status:	Completed
Principal investigator:	William Underwood (technologist), Georgia Tech Research Institute
Study team members:	Kevin Glick (archivist), Yale University Mark Wolfe (archivist), State University of New York, Albany

APPENDIX 4

Genesis of the Case Study Research Questions

Original Questions From Workshop #2 (21 June 2002, Wash., D.C.)	**Re-worded Questions Distributed by Project Director (26 June 2002)**	**Additional Questions Added During Workshop #3 (17-21 Sept 2002, Los Angeles)**
	1. What activities of the creator are you investigating?	
	2. Which of these activities generate the digital entities that are the objects of your case study?	
1. What are the purposes of the information you record or create?	4. For what purpose(s) are the digital entities you are examining created?	
2. What information do you create to meet these purposes?		
3. What methods and paradigms inform your work?		
4. What forms do your information take?	3. What form do these digital entities take? (E.g. e-mail, CAD, database)	
5. What processes do you follow in creating information?	5. How are those digital entities generated?	
6. What are the key processes in creating the information?	6. From what precise process(es) or procedure(s), or part thereof, do the digital entities emerge?	
	7. To what other digital or non-digital entities are they connected? Is such connection documented or captured?	
7. How do you record and identify the information, the methods and the technologies you have followed?	8. What are the documentary and technological processes or procedures that the creator follows to identify, retrieve, and access the digital entities?	

8. How do you document the processes and procedures you use?	9. Are those processes and procedures documented? How? In what form?	
9. What are the key elements of the information you create?	10. What are the key elements, attributes, and digital components of the entities under examination? Should be asked of the interviewee (opinion question, subject to later analysis).	
10. What measures do you take to ensure the quality and reliability of the information you create or information sources that you use?	11. What measures does the creator take to ensure the quality, reliability and authenticity of the digital entities and their documentation? If no specific measure is taken, does the creator think that those qualities are to be assumed for its digital entities? *(Note overlap with question 16.)*	
12. How are the changes made to your information and how are these recorded?	12. How are changes to the digital entities made and recorded?	
11. How do you use the information you create?	13. How does the creator use the digital entities under examination?	
13. How do others use the information you create?	14. Do external users have access to the digital entities in question? If so, how, and what kind of uses are made of the entities?	
14. Do others add to your information to create new information?		
15. What do you conceive of as a record?	15. Among its digital entities, which ones does the creator consider to be records and why?	

16. What do you conceive of as an authentic record?	16. Does the creator think that the authenticity of his digital records is assured, and if so, why?	
17. How do you preserve this authentic record?	17. Does the creator keep the digital entities that are currently being examined? That is, are these digital entities part of a record keeping system? If so, what are its features?	
18. How do you preserve this through technological change?	18. How does the creator maintain its digital entities through technological change?	
		19. Have you had to make rules, or adopt standards to help you in your work? Do you find you have to update them regularly?
		20. Do any legal or ethical issues arise from your electronic work?
		21. Did you create or adopt a standard list of information which you try to record about each file, or work?
		22. Where did you get it? Do you know if others use the same one?

APPENDIX 5

23 Case Study Questions that the researchers should be able to answer at the completion of their investigation

23 Case Study Questions that the researchers should be able to answer at the completion of their investigation

March 18, 2003

1. What activities of the creator have you investigated?

2. Which of these activities generate the digital entities that are the objects of your case study?

3. For what purpose(s) are the digital entities you have examined created?

4. What form do these digital entities take? (e.g. e-mail, CAD, database)
   - 4a. What are the key formal elements, attributes, and behaviour (if any) of the digital entities?
   - 4b. What are the digital components of which they consist and their specifications?
   - 4c. What is the relationship between the intellectual aspects and the technical components?
   - 4d. How are the digital entities identified (e.g., is there a [persistent] unique identifier)?
   - 4e. In the organization of the digital entities, what kind of aggregation levels exist, if any?
   - 4f. What determines the way in which the digital entities are organized?

5. How are those digital entities created?
   - 5a. What is the nature of the system(s) with which they are created? (e.g. functionality, software, hardware, peripherals etc.)
   - 5b. Does the system manage the complete range of digital entities created in the identified activity or activities for the organization (or part of it) in which they operate?

6. From what precise process(es) or procedure(s), or part thereof, do the digital entities result?

7. To what other digital or non-digital entities are they connected in either a conceptual or a technical way? Is such connection documented or captured?

8. What are the documentary and technological processes or procedures that the creator follows to identify, retrieve, and access the digital entities?

9. Are those processes and procedures documented? How? In what form?

10. What measures does the creator take to ensure the quality, reliability and authenticity of the digital entities and their documentation?

11. Does the creator think that the authenticity of his digital entities is assured, and if so, why?

12. How does the creator use the digital entities under examination?

13. How are changes to the digital entities made and recorded?

14. Do external users have access to the digital entities in question? If so, how, and what kind of uses do they make of the entities?

15. Are there specific job competencies (or responsibilities) with respect to the creation, maintenance, and/or use of the digital entities? If yes, what are they?

16. Are the access rights (to objects and/or systems) connected to the job competence of the responsible person? If yes, what are they?

17. Among its digital entities, which ones does the creator consider to be records and why?

18. Does the creator keep the digital entities that are currently being examined? That is, are these digital entities part of a record keeping system? If so, what are its features?
    - 18a. Do the recordkeeping system(s) (or processes) routinely capture all digital entities within the scope of the activity it covers?
    - 18b. From what applications do the recordkeeping system(s) inherit or capture the digital entities and the related metadata (e.g. email, tracking systems, workflow systems, office systems, databases, etc.)?
    - 18c. Are the digital entities organized in a way that reflects the creation processes? What is the schema, if any, for organizing the digital entities?
    - 18d. Does the recordkeeping system provide ready access to all relevant digital entities and related metadata?
    - 18e. Does the recordkeeping system document all actions/transactions that take place in the system re: the digital entities? If so, what are the metadata captured?

19. How does the creator maintain its digital entities through technological change?
    - 19a. What preservation strategies and/or methods are implemented and how?
    - 19b. Are these strategies or methods determined by the type of digital entities (in a technical sense) or by other criteria? If the latter, what criteria?

20. To what extent do policies, procedures, and standards currently control records creation, maintenance, preservation and use in the context of the creator's activity? Do these policies, procedures, and standards need to be modified or augmented?

21. What legal, moral (e.g. control over artistic expression) or ethical obligations, concerns or issues exist regarding the creation, maintenance, preservation and use of the records in the context of the creator's activity?

22. What descriptive or other metadata schema or standards are currently being used in the creation, maintenance, use and preservation of the recordkeeping system or environment being studied?

23. What is the source of these descriptive or other metadata schema or standards (institutional convention, professional body, international standard, individual practice, etc.?)

APPENDIX 6

Call for Case Study Proposals

Call for Case Study Proposals InterPARES 2 April 2002

Overview

Beginning in 2002, researchers in the InterPARES 2 Project will undertake a set of case studies in an effort to address in part the research questions assigned to each domain task force and cross-domain research team. Each case study will focus on the records (or some portion of the records) and records management process of a specific creator. Within the context of a case study, it may also be possible to test the appraisal and preservation models developed in InterPARES 1 and/or develop a prototypical system for the preservation of the records in question.

In March 2002, many InterPARES researchers participated in an exercise involving the sketching of records creation processes related to an activity with which they were familiar. This exercise has proven fruitful in identifying potential case study subjects and specific issues of concern. A number of case studies are currently in development based on these activity models, with more to follow in the coming months.

At the International Team Meeting in June 2002, the chairs of the various working groups and cross-domain teams will review the case study proposals that have been developed, offer suggestions to make the conduct of each case study more efficient and effective, and allocate resources as needed. Thus, the goal of the review as envisaged will not be to approve or reject case study proposals, but rather to coordinate the efforts of researchers, allow for the streamlining and sharing of research tools and offer advice and support to researchers conducting case studies.

This call is designed to assist researchers in the development of case study proposals. It is expected that researchers interested in proposing and/or leading a case study will be in contact with the chairs of their working groups as well as the InterPARES administration as their case study is developing. The International Team will review case study proposals as they are ready; those wishing their proposal to be reviewed at the meeting in June should submit their proposals by June 10 2002 to allow for distribution prior to the meeting.

InterPARES 1 Case Studies

The Authenticity Task Force (ATF) of the InterPARES 1 Project undertook more than thirty case studies (in four rounds) of a variety of electronic systems. The initial goal of the ATF was to identify the formal elements shared by all electronic records, the elements that allowed for their differentiation by type, and the elements that allowed a record's authenticity to be verified over time. (It was further hypothesized that answers to these questions would lead to conclusions regarding the possibility of migrating electronic records from one system to another without compromising their authenticity.) Such analysis of a record's formal elements is *diplomatic* analysis; diplomatics as a science encompasses a set of principles and terminology that have been used to analyze records since the 17th century. Given that electronic records in most cases serve the same administrative functions that paper records have in the past, InterPARES 1 researchers were looking for possible parallels between paper and electronic records, specifically in their formal elements. To this end, the ATF developed a template for analysis of electronic records, enumerating and describing a large set of formal elements that might be potentially found in an electronic record. Once the template had been finalized, a number of case studies of electronic systems were undertaken to empirically ascertain whether or not these elements manifested themselves in records contained within actual electronic systems.

To gather information about actual electronic systems and records, the ATF developed two tools: a Case Study Interview Protocol (CSIP) and a Template Element Data Gathering Instrument (TEDGI). The CSIP was a set of questions posed by a researcher to individuals familiar with the workings of the electronic system being studied. Based on the interviewees' responses to the CSIP questions, the researcher noted (in the TEDGI) the presence or absence of formal record elements. This two-step process was deemed necessary in light of the interviewees' unfamiliarity with diplomatic terminology. The ATF reported that the case study analysis did indeed confirm the presence or absence of certain formal elements and further indicated potential weaknesses in many of the records management system studied. It was also observed that many of the formal elements upon which a record's authenticity was presumed, which in the past had been visibly manifested on the face of the record, were supplanted in many electronic systems by procedural and technological controls.

In general the ATF concluded that the case study approach was very useful in addressing their assigned research questions, but had certain limitations and could be improved upon. The ATF recommended that in future case study research that 1) the record-keeping system be studied as a whole (including its paper elements), 2) the study begin with a careful analysis of the business procedures of the record creators in question in order to identify the actions in which records participate, 3) formulate a questionnaire (or revise and shorten the CSIP) with terminology familiar to the interviewees and 4) delve further into the technological context of the records under study. (Footnote 1: see the Authenticity Task Force Report on the InterPARES Web site http://www.interpares.org/reports.htm.)

It is hoped that in InterPARES 2 we can implement these methodological findings of InterPARES 1. Please keep these findings in mind when developing case studies and case study tools.

The Case Study Proposal

Case study proposals will include a description of the case study subject, the rationale for choosing that case study subject, the research methodologies to be employed, a description of the research team and their roles, and a timeline.

- Description of the case study subject: briefly describe the case study subject in terms of their business mandate and business processes. If the records of interest to us are only a portion of the creator's records, contextualize the records within the creator's records as a whole.
- Rationale: why are this creator and these records of interest to InterPARES 2? Which research questions will be addressed in this case study?
- Research methodologies: how will the case study be conducted? What sorts of data will be gathered? How will the data be represented? What tools will be used in gathering the data? Below are listed a handful of possible methodologies and related issues:
    a. Interviewing records creators: much information can be gathered by interviewing the creators of records. The questionnaire used in InterPARES 1 and the research questions for InterPARES 2 form a starting point for developing a set of questions; further questions may be appropriate for specific creators.

b. Modeling records creation processes: collaboratively creating a model of the creator's business process may illuminate issues unanticipated in a questionnaire. Where applicable, the process models sketched in March, 2002 may serve as an outline. Also, a formal work-flow model may already exist for a given organization.
c. Application of IP1 Appraisal model: as much of the information that needs to be gathered is similar to the information typically gathered by archivists conducting appraisals, it may be a useful exercise to use the InterPARES 1 appraisal model to guide the study.
d. Replicating systems for hands-on study of records: though difficult with large systems, creators working on desk-top computers may be willing to make copies of all or a portion of their records which can be installed on an InterPARES computer for subsequent study. It may also be possible to replicate a larger system (such as a government Web site) and create a set of records for testing. This methodology would greatly facilitate diplomatic analysis of the records in questions, as they could be examined in their native environment.
e. Testing preservation strategies: it may be advantageous to transfer a sampling of records to an InterPARES computer to test various preservation strategies, or to walk through the InterPARES 1 preservation model. It may also be beneficial to prototype a preservation system for the records in question.

- Case study team: each case study will be conducted by a team of researchers. This team will comprise a lead investigator, other interested researchers, and research assistants as necessary.
- Timeline: a timeline will outline the sequence and timing of activities in the conduct of the case study, as well as the researchers responsible for those activities (individually or collectively). Keep in mind that the team will have a chance to meet at the workshops in September 2002 and February 2003; further meetings may be conducted electronically. If further face-to-face meetings are necessary and require travel, include a estimate of the costs involved. In developing the timeline, consider the time and effort necessary to make arrangements with the case study subject, to develop a questionnaire, model or prototype, to seek human subjects approval, to represent and analyze the data, to follow up with the case study subject (pursuant questions, clarification of responses, validation of a process model, etc.) and to prepare a case study report.

Reporting Procedure

Given the flexibility in the conduct of case studies, it is important that the reporting procedure be harmonized among case studies so as to facilitate comparison. The case study report will begin with information included in the proposal (description of subject, rationale, methodologies used), noting any modifications made in actually conducting the case study. The report will also include the case study data or summary of same, the observations and comments of the case study team and an evaluation of the effectiveness of the methodologies employed. Ultimately, case study reports will be reviewed by the appropriate task forces and will serve as the basis of much of their deliberations.

APPENDIX 7

Diplomatic Analysis Template

Diplomatic Analysis Template

Diplomatic Analysis

CS[##]
[Title of Case Study] Case Study

INTRODUCTION

Paragraph 1[1] - The InterPARES case study # …

Paragraph 2[2] - The … project/program/system/database …

Paragraph 3[3] - The following text presents the results of the diplomatic analysis on the digital entity identified in the case study report. The purpose of the diplomatic analysis is to assess the status of the identified digital entity as record, and based on the analysis, Domain 3 of InterPARES 2 could propose applicable preservation strategies. The digital entity identified in the case study report is … This diplomatic analysis therefore centers on the identification of the [digital entity] as record.

IDENTIFICATION OF RECORD(S)

The current version of the InterPARES glossary definition of record defines a record[4] as "a document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference."[5] This definition implies that, to be considered as a record, a digital entity must comprise five indispensable elements: fixed content and form, embedded action, archival bond, persons and contexts, the establishment of which is based on diplomatic analysis, archival science and findings from InterPARES. The application of the definition on the [digital entity] is therefore analyzed accordingly:

1. **TO BE IDENTIFIED AS A RECORD, THE DIGITAL ENTITY MUST POSSESS FIXED CONTENT AND FORM,[6] AND BE AFFIXED TO A STABLE MEDIUM (OR PHYSICAL CARRIER).**
   - The content of the [digital entity] is fixed or not and why
   - The documentary form[7] of the [digital entity] is fixed or not and why

---

[1] Introduction to the case study.
[2] Introduction to the case study subject.
[3] Introduction to the diplomatic analysis, including the indication of the subject of the diplomatic analysis, i.e., the digital entity in the report that needs to be analyzed.
[4] Current version of the definition of record as assessing criterion, which starts the process of analyzing the 5 components.
[5] Glossary definitions, in Terminology Database, accessible through the InterPARES website research restricted area.
[6] The InterPARES1 Authenticity Task Force has defined fixed form as the following: 1) binary content of the record, including indicators of documentary form, must be stored in a manner that ensures it remains complete and unaltered, and 2) technology must be maintained and procedures defined and enforced to ensure that the content is presented or rendered with the same documentary form it had when set aside. (See ATF Research Methodology Statement, available at: http://www.interpares.org/documents/interpares_ResearchMethodologyStatement.pdf).
[7] Definition of documentary form from Glossary Definitions, Terminology Database, InterPARES Web site: The rules of representation according to which the content of a record, its administrative and documentary context, and its authority are communicated. Documentary form possesses both extrinsic and intrinsic elements.

- The digital entity is fixed to a stable medium or not and why

2. **A RECORD MUST ALSO PARTICIPATE IN AN ACTION, DEFINED AS THE CONSCIOUS EXERCISE OF WILL BY THE AUTHOR OR BY AN EXTERNAL PERSON, AIMED TO CREATE, MAINTAIN, MODIFY OR EXTINGUISH SITUATIONS. A RECORD RESULTS AS A NATURAL BY-PRODUCT OF THE ACTION.**

   Statement of the name of action that generated the [digital entity] and how it participates in the action.

3. **THIRD, A RECORD MUST POSSESS AN ARCHIVAL BOND, WHICH IS THE RELATIONSHIP THAT LINKS EACH RECORD TO THE PREVIOUS AND SUBSEQUENT RECORD OF THE SAME ACTION AND, INCREMENTALLY, TO ALL THE RECORDS WHICH PARTICIPATE IN THE SAME ACTIVITY. THE ARCHIVAL BOND IS ORIGINARY (I.E., IT COMES INTO EXISTENCE WHEN THE RECORD IS MADE OR RECEIVED AND SET ASIDE), NECESSARY (I.E., IT EXISTS FOR EVERY RECORD), AND DETERMINED (I.E., IT IS CHARACTERISED BY THE PURPOSE OF THE RECORD).**

   The [digital entity] possesses (or does not possess) archival bond and why.

4. **FOURTH, RECORD CREATION MUST INVOLVE AT LEAST THREE PERSONS, WHETHER OR NOT THEY EXPLICITLY APPEAR IN THE RECORD ITSELF. THESE PERSONS ARE THE AUTHOR, ADDRESSEE AND WRITER; IN THE ELECTRONIC ENVIRONMENT, ONE MUST ALSO TAKE INTO ACCOUNT TWO ADDITIONAL NECESSARY PERSONS: THE CREATOR AND THE ORIGINATOR.**

   - The record's author is the physical or juridical person having the authority and capacity to issue the record or in whose name or by whose command the record has been issued.

     Name of the author and why.

   - The writer is the physical or juridical person having the authority and capacity to articulate the content of the record.

     Name of the writer and why.

   - The addressee is the physical or juridical person(s) to whom the record is directed or for whom the record is intended.

     Name of the address and why.

   - The creator is the person in whose fonds the record exists.

     Name of the creator and why.

   - The originator is the person to whom the Internet account issuing or the server holding the record belongs.

     Name of the originator and why.

5. FINALLY, A RECORD MUST POSSESS AN IDENTIFIABLE CONTEXT, DEFINED AS THE FRAMEWORK IN WHICH THE ACTION IN WHICH THE RECORD PARTICIPATES TAKES PLACE. THE TYPES OF CONTEXT INCLUDE JURIDICAL-ADMINISTRATIVE, PROVENANCIAL, PROCEDURAL, DOCUMENTARY, AND TECHNOLOGICAL.

- The juridical-administrative context is the legal and organizational system in which the creating body belongs.

  Juridical context:
  Administrative context:

- The provenancial context refers to the creating body, its mandate, structure and functions.

  Provenancial context:

- The procedural context comprises of the business procedure in the course of which the record is created.

  Procedures:
  Diplomatic analysis of Procedural Phases in the Creation of the [digital entity]

  a) ***Initiative*****:** The introductory phase of any procedure is "constituted by those acts, written and/ or oral, which start the mechanism of the procedure."[8]

  b) ***Inquiry***: This preliminary phase "is constituted by the collection of the elements necessary to evaluate the situation."[9]

  c) ***Consultation*****:** This phase is "constituted by the collection of opinions and advice after all the relevant data has been assembled."[10]

  d) ***Deliberation:*** This phase is "constituted by the final decision-making."[11]

  e) ***Deliberation control*****:** This phase is "constituted by the control exercised by a physical or juridical person different from the author of the document embodying the transaction, on the substance of the deliberation and / or on its forms."[12]

  f) ***Execution*****:** "The documents created in this phase are the originals of those embodying the transactions."[13] In other words, the execution phase results in the issuing of the first record capable of producing the consequences intended by its author.

---

[8] Luciana Duranti, *Diplomatics: New Uses for an Old Science* (Lanham, Maryland and London: The Scarecrow Press in association with the Society of American Archivists and the Association of Canadian Archivists, 1998), 115.
[9] Ibid.
[10] Ibid.
[11] Ibid.
[12] Ibid.
[13] Ibid., 116.

- The documentary context is defined as the archival fonds to which a record belongs and its internal structure.

  Documentary context:

- The technological context is defined as the characteristics of the technological components of an electronic computing system in which records are created.

  Technological context:

CONCLUSIONS

Overall conclusions of the status of the digital entity under examination:

If it is not a record:

Summary of the digital entity as a publication:
Summary of the digital entity as a potential record:

APPENDIX 8

InterPARES 2 Reporting Framework

InterPARES 2 Reporting Framework
December 2003

Case study reports should contain the following sections:

A. Overview
B. Statement of Methodology
C. Description of Context
D. Narrative answers to the core research questions
E. Narrative answers to applicable domain and cross-domain research questions
F. Bibliography of relevant material, including articles about the methods and works of the subject(s)
G. Glossary of terms
H. Preliminary model

A. Give a brief overview of the subject and the nature of the case study.

B. Briefly describe the data gathering and analysis methodology employed to achieve the research objective of answering the core twenty-three questions posed to researchers, and any applicable domain and cross-domain questions.

For most or all artistic focus case studies the primary information-gathering tool will be an interview or interviews that draws upon the lay restating and amplification of these questions (ip2_possible%20cs_interview_questions.pdf). Note that while recording and creating transcripts of interviews is highly recommended, transcripts should not be included in the case study report.

Other research methods might include document review, ethnographic analysis or participant observation, diplomatic analysis, Bayesian analysis, content analysis (of interview transcripts), etc.

C. Describe the context of record or digital entity creation and management. InterPARES 1 defined five contexts.

a. Provenancial: the creating body, its mandate, structure, and functions (indicators include organizational charts, annual reports, the classification scheme, etc.).

b. Juridical-administrative: the legal and organizational system in which the creating body (indicated by laws, regulations, etc.) belongs.

c. Procedural: the business procedure in the course of which the digital entity is created (indicators include workflow rules, codes of administrative procedure, classification schemes, etc.).

d. Documentary: the fonds to which the digital entity belongs and its internal structure (indicators include classification schemes, record inventories, indexes, registers, etc.) *Note: In some organizations, business procedures are integrated with documentary procedures.*

e. Technological: the characteristics of the digital environment in which the record is created and maintained.

These specifications assume a legal and bureaucratic structure that might not be appropriate to all case studies. However, each of these contexts should be addressed at least to the extent that their inapplicability is documented, and the environment in which digital entities are created and managed, or the framework of action in which they participate, should be described.

D. Narrative answers to the twenty-three core questions to researchers listed in ip2_23_questions.pdf.

E. Narrative answers to any relevant domain and cross-domain research questions.

F. The bibliography should draw from those bibliographies already created where appropriate, and additionally list any articles or monographs that are of particular relevance to the specific case study.

G. The glossary should list and define the key terms used in the case study, both for purposes of possible inclusion in the IP2 glossary, and to allow definitions to be compared with those that already exist within the IP2 glossary.

H. A preliminary case study model should be made following the guidelines provided by the modeling cross-domain group.

APPENDIX 9

Domain 1 Research Questions

Domain 1 Research Questions

1a) What types of documents are traditionally created (that is, made or received) and set aside in the course of these activities that are expected to be delivered online? For what purposes?

1b) What types of electronic documents are currently being created to accomplish those same activities? Have the purposes for which these documents are created changed?

2) What are the nature and the characteristics of the traditional process of document creation in each activity? Have they been altered by the use of digital technology and, if yes, how?

3) What are the formal elements and attributes of the documents generated by these processes in both a traditional and a digital environment? What is the function of each element and the significance of each attribute? Specifically, what is the manifestation of authorship in the records of each activity and its implications for the exercise of intellectual property rights and the attribution of responsibilities?

4) Does the definition of a record adopted by InterPARES 1 apply to all or part of the documents generated by these processes? If yes, given the different manifestations of the record's nature in such documents, how do we recognize and demonstrate the necessary components that the definition identifies? If no, is it possible to change the definition maintaining theoretical consistency in the identification of documents as records across the spectrum of human activities? In other words, should we be looking at factors that make a document a record other than those that diplomatics and archival science have considered so far?

5) As government and businesses deliver services electronically and enter into transactions based on more dynamic Web-based presentations and exchanges of information, are they neglecting to capture adequate documentary evidence of the occurrence of these transactions?

6) Is the move to more dynamic and open-ended exchanges of information blurring the responsibilities and altering the legal liabilities of the participants in electronic transactions?

7) How do record creators traditionally determine the retention of their records and implement this determination in the context of each activity? How do record retention decisions and practices differ for individual and institutional creators? How has the use of digital technology affected their decisions and practices?

APPENDIX 10

Template for Case Study Analysis ("Areas to be Covered")

by

Heather Daly and Ann Forman
School of Library, Archival and Information Studies,
The University of British Columbia

Template for Case Study Analysis ("Areas to be Covered")

Version 1.0: Heather Daly and Ann Forman, December 2004
(Pagination may differ from original document: ip2(d1)_anlaysis_template.pdf)

The purpose of this template is to enable the gathering of information spread over the whole of the documentation related to each case study. Through this exercise, a general overview will be created, which will provide insight into the types of record creators and activities that have been studied by Focus 1, 2 and 3. With this overview, we will be able to validate the case studies and inform the work of the Domains and Cross-domains.

This template is not intended to be another framework to be used when writing the Final Report. Rather, it serves as a reminder of the information that we must have as part of the Report.

The template has been structured in two sections. The first concerns the ***Creator of the Records***. The second concerns the ***Activities Resulting in Document Creation***. This latter section is split into two sub-sections, involving the "Administrative and Managerial Framework" and the "Digital Entity/Entities Under Study."

CREATOR OF THE RECORDS

The information we are looking for about the *creator of the records* is often embedded within the provenancial, juridical-administrative, procedural, documentary and technological contexts (see Case Study Reporting Framework, December 2003). This information is essential to understanding who has produced the digital records, and for what reasons. This understanding will allow the characterization of the case studies.

If the following is inapplicable to the case study at hand or unavailable, please indicate wherever possible in the report why this is so.

Name

- Provide the official name and other names of the body under study.

Location

- Provide the country, region (example: province, state) and/or city which exerts the most legal influence over the body.

Origins

- Provide the origins of the body, such as information regarding how and why the body began its activities.
- Provide the official founding date and/or founding event.

Legal Status

- Provide the legal status. For example: “private individual,” “for-profit small company,” “research group”
- Provide the year of legal establishment, if applicable.
- Provide specifics about the most relevant laws under which body is governed. For example: “copyright legislation,” “Companies Act”
- Provide information about any legal status inherited from other organizations or associations, any other legally required standards, codes or regulations that apply to the body.

Norms

- Provide information about any non-legally required standards, methodologies, codes or regulations that are subscribed to by the body.
- Provide information about the non-legally required standards, methodologies, codes or regulations from other organizations, traditions or associations that are subscribed to by the body. For example: “methodologies related to archaeology”

Funding

- Provide information about the sources of revenue related to the digital entity under study. For example: “grants”, “ticket sales”

Resources (Physical)

- Provide information about the physical context in which the creator is working, including relevant information about equipment and infrastructure. For example: “one office, shared with another group”

Governance

- Provide information about how the body is managed. For example: “cooperative,” “collective,” “partnership”
- Provide information about the organization of the body, such as through the inclusion of an organizational chart.
  Provide information about employees, members or partners (number, areas of specialization, qualifications, turnover).
- Provide information about the body’s place within an organization, if applicable.
- Provide information about any internal policies or regulations.

Mandate

The responsibilities of the body

- Provide information about the responsibilities of the body given to the body through enabling legislation.
- Provide information about any stated mandate.

Philosophy

The vision and values the body works toward or under

- Provide information about the body’s philosophy and values.

- Provide information about which genres or disciplines the body is related. For example: “multidisciplinary theatre,” “geology”
- Provide information about the schools of thought to which the body subscribes, if these influence the body’s choices and practices.

Mission

The stated ways in which the body is working towards the mandate

- Provide the mission statement(s), which may have evolved over time.

Functions

- List all of the major functions which the body undertakes to fulfil the mission(s) and mandate. For example: “administration,” “research,” “performance,” “training”

Recognitions

- Provide information about any achievements, honours or prizes that the body has received for its work.

ACTIVITIES RESULTING IN DOCUMENT CREATION

Administrative and Managerial Framework

This section is divided into two sub-sections. The first concerns the Administrative and Managerial Framework within which the digital entity under study is created. The second focuses on the digital entity. Both sub-sections aim to gather information to allow the characterization of the types of activities and entities that have been studied.

If the following is inapplicable to the case study at hand or unavailable, please indicate wherever possible in the report why this is so.

General description

- Provide a general description of administrative practices. For example: “The creator must administer payroll and grant applications”

Type of activities

- List the general types of administrative activities undertaken on a regular basis. For example: “submitting reports,” “writing grant applications”

Documents resulting from activities

- List the main types of documents resulting from administrative activities. For example: “receipts,” “reports,” “correspondence”

Existence of a records management and/or archives program

- Provide a description of the existence of activities related to archives and records management.
- Provide information about any policies that the body may have which govern archives and records management.

Individuals responsible for preservation

- Provide the name and qualifications of individuals(s) responsible for archives and/or records management.
- Provide information about the relationship of the individuals responsible for preservation to the creation of the records. For example: "Once completed, he maintains the records on his computer"

Existence of preservation strategies

- Provide the location in which the records are kept.
- Provide the nature in which records are kept. For example: "All records are digitized"
- Provide a description of the organization of the records created by the body under study. For example: "Records are split, with some records being kept by the contracting party"
- Provide a brief description of any methods used to preserve records.
- Provide a brief description of any methods used to attempt to avoid technological obsolescence.

Legal Requirements and Constraints

- Provide a description of how the relevant laws impact upon the policies and procedures by which *administrative activities* are carried out.
- Provide a description of how the relevant laws impact upon the creation, form, content, identity integrity, organization and preservation of the records related to administrative activities.

Normative Requirements and Constraints

The written or unwritten rules of a specific discipline or area of thought to which the body subscribes

The written or unwritten rules may not be limited to scientific, artistic and ethical requirements and constraints.

<u>Scientific requirements and constraints</u>

Scientific foundations of the discipline with which the body uses or identifies with that require, influence or prohibit certain behaviours

- Provide a description of how relevant scientific requirements/ constraints impact upon the policies and procedures by which administrative activities are carried out.
- Provide a description of how relevant scientific requirements/ constraints impact upon the creation, form, content, identity, integrity, organization and preservation of the records related to administrative activities.

<u>Artistic requirements and constraints</u>

Artistic foundations or schools of thought which the body uses or identifies with that require, influence or prohibit certain behaviours

- Provide a description of how relevant artistic requirements/constraints impact upon the policies and procedures by which administrative activities are carried out.
- Provide a description of how relevant artistic requirements/constraints impact upon the creation, form, content, identity, integrity, organization and preservation of the records related to administrative activities.

Ethical requirements and constraints

Propriety and rules of behaviour which the body uses or identifies with that require, influence or prohibit certain behaviours

- Provide a description of how relevant ethical requirements/constraints impact upon the policies and procedures by which administrative activities are carried out.
- Provide a description of how relevant ethical requirements/constraints impact upon the creation, form, content, identity, integrity, organization and preservation of the records related to administrative activities.

Technological Requirements and Constraints

Technology requirements and constraints related only to the administrative or management function

- Provide a description of the equipment used:
    - Architecture (e.g., network topology, infrastructure, hardware)
    - Creation or input tools (e.g., software, camera, microphone)
    - Processing tools (e.g., software, console)
- Provide a list of the types of media created (e.g., graphic, textual, audio).
- Provide a list of the formats created (e.g., .pdf, .doc, .jpg).
- Provide a description of how relevant technological requirements/constraints impact upon on the policies and procedures by which administrative activities are carried out.
- Provide a description of how relevant technological requirements/constraints impact upon the creation, form, content, identity integrity, organization and preservation of the records related to administrative activities.

ACTIVITIES RESULTING IN DOCUMENT CREATION

Digital Entity/Entities Under Study

General description of the activity

- Provide the name and type of the digital entity/entities being studied.
- Provide a description of the goals and functions of the digital entity/entities.
- Provide a description of how the digital entity/entities relate(s) to the body's mandate and mission(s).

Type of activities

- Provide a description of the activities related to the creation of the digital entity/entities.

Documents resulting from activities

- Provide a list of documents that enable the activities related to the digital entity/entities.. For example: "photographs," "correspondence"
- Provide a list of documents that result from the activities related to the digital entity/entities. For example: "digitized images," "e-mails"

Existence of preservation strategies

- Provide a brief description of any methods used to preserve records related to the digital entity/entities or the digital entity/entities itself/themselves.

- Provide a brief description of any methods used to attempt to avoid technological obsolescence.

Legal Requirements and Constraints

- Provide a description of how any relevant laws impact upon the policies and procedures by which activities related to digital entity/entities are carried out.
- Provide a description of how any relevant laws impact upon the creation, form, content, identity, integrity, organization and preservation of the records generated by the digital entity/entities or the digital entity/entities of the record(s).

Normative Requirements and Constraints

The written or unwritten rules of a specific discipline or area of thought to which the body subscribes

The written or unwritten rules may not be limited to scientific, artistic and ethical requirements and constraints.

<u>Scientific requirements and constraints</u>

Scientific foundations of the discipline with which the body uses or identifies with that require, influence or prohibit certain behaviours

- Provide a description of how relevant scientific requirements/constraints impact upon the policies and procedures by which activities related to digital entity/entities are carried out.
- Provide a description of how relevant scientific requirements/constraints impact upon the creation, form, content, identity, integrity, organization and preservation of the records generated by the digital entity/entities or the digital entity/entities of the record(s).

<u>Artistic requirements and constraints</u>

Artistic foundations or schools of thought which the body uses or identifies with that require, influence or prohibit certain behaviours

- Provide a description of how relevant artistic requirements/constraints impact upon the policies and procedures by which activities related to digital entity/entities are carried out.
- Provide a description of how relevant artistic requirements/ constraints impact upon the creation, form, content, identity, integrity, organization and preservation of the records generated by the digital entity/entities or the digital entity/entities of the record(s).

<u>Ethical requirements and constraints</u>

Propriety and rules of behaviour which the body uses or identifies with that require, influence or prohibit certain behaviours

- Provide a description of how relevant ethical requirements/constraints impact upon the policies and procedures by which activities related to digital entity/entities are carried out.
- Provide a description of how relevant ethical requirements/constraints impact upon the creation, form, content, identity, integrity, organization and preservation of the records generated by the digital entity/entities or the digital entity/entities of the record(s).

Technological Requirements and Constraints

Technology requirements and constraints related only to the administrative or management function

- Provide a description of the equipment used:
    - Architecture (e.g., network topology, infrastructure, hardware)
    - Creation or input tools (e.g., software, camera, microphone)
    - Processing tools (e.g., software, console)
- Provide a list of the types of media created (e.g., graphic, textual, audio).
- Provide a list of the formats created (e.g., .pdf, .doc, .jpg).
- Provide a description of how relevant technological requirements/constraints impact upon on the policies and procedures by which activities related to digital entity/entities are carried out.
- Provide a description of how relevant technological requirements/constraints impact upon the creation, form, content, identity, integrity, organization and preservation of the records related to the digital entity/entities of the record(s).

APPENDIX 11

Case Studies at-a-Glance

by

Geneviève Shepherd
School of Library, Archival and Information Studies,
The University of British Columbia

Case Studies at-a-Glance

Geneviève Shepherd, The University of British Columbia
10 Mar 2006, updated 16 Apr 2008

Case Study and General Study Statistics

Completed Case Studies
Number of Case Studies completed to date: **23***
ARTS: **10**
SCIENCE: **5**
GOVERNMENT: **8**

Number Analyzed and Characterized: **23**
Number of Analyses Validated: **23**
Number of Characterizations Validated: **23**
**Includes four sections of CS09; excludes CS22 draft final report completed by UBC GRAs*

Case Studies Yet To Be Completed and ***Interim Reports***
Number of Case Studies yet to be completed: **1**
ARTS: **1***
SCIENCE: **0**
GOVERNMENT: **0**
Number Analyzed and Characterized: **1****
Number of Analyses Validated: **0**
Number of Characterizations Validated: **0**
**This is CS22, for which there is a yet-to-be-validated draft report that has been used to complete the analysis and characterization. Note: This case study has since been retired.*
****CS22, based on draft report*

General Studies
Number of General Studies completed to date: **10**
Number of General Studies yet to be completed: **1***
Number of General Studies Analyzed and Characterized: **0**
**This is GS02. Note: This general study has since been retired.*

Case Study Documentation Summary

	01	02	03	05	06	08	09/1	09/2	09/3	09/4	10	12	13	14	15	17	18	19	20	21	22	24	25	26
Interim/Draft Report(s)	X				X						X		X	X							X[1]	X		X
Case Study (Final) Report	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X
• **Glossary of Terms/Abbreviations**	X	X		X		X						X	X		X			X	X			X		X
• **Workflow Diagram**	X	X	X	X	X	X	X			X	X													
• **Bibliography** (in Report)	X	X			X	X					X	X	X			X	X	X	X	X			X	
• **Other**[2]				X	X	X	X		X			X		X		X	X		X	X		X		
Areas to Be Covered (CS)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Characterization (CH)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
CS/CH Validation	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X
Diplomatic Analysis	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Entity Relationship Model[3]	X	X											X	X	X									
Activity Model[3]		X		X			X					X	X	X	X		X	X	X	X		X		X
Bibliography, Posted Separately		X			X								X											
Literature Review				X[4]	X																			
Survey/Questionnaire Results					X									X[5]										
Transcripts/Interview Notes	X	X	X		X									X	X									
Presentation(s)[6]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X		X
Article(s)					X		X	X	X	X		X		X				X				X		
Research Documentation[7]	X			X	X	X			X					X			X							

[1] Report completed by GRAs, but never validated by the Principle Investigator.
[2] e.g., list of laws, abstracts, productions; screenshots; charts, etc.
[3] Copy may be in final report.
[4] A summary of the literature survey is posted; the full literature survey is part of the Report (Appendix).
[5] The survey has since been extracted from the case study and designated as general study 09.
[6] e.g., PowerPoint.
[7] e.g., ethics form, methodology, etc.

Case Studies at-a-Glance

The following information is based on case study proposals and validated and yet-to-be validated case study analyses and characterizations.

CS01: Arbo Cyber théâtre (?)
Final Report: Yes Focus: Arts
Type of Discipline: Theatre
Type of Creator: Theatre group
Type of Organization: Private corporation
Managerial Framework: Two individuals in Québec City, QC, Canada
Type of Digital Entity: Web site (*Ludosynthèse*). Purpose: maintain memory of group, while allowing audience interaction to continue

CS02: Performance Artist Stelarc
Final Report: Yes Focus: Arts
Type of Discipline: Performance art
Type of Creator: Private individual
Type of Organization: Individual; may work anywhere but is from Australia
Managerial Framework: Individual
Type of Digital Entity: Web site. Purpose: advertising, and implementing and documenting the stages of the performance process

CS03: *HorizonZero*/ZeroHorizon Online Magazine and Media Database
Final Report: Yes Focus: Arts
Type of Discipline: Media
Type of Creator: Media and Visual Arts Department
Type of Organization: Institute, part of larger Centre; in Banff, AB, Canada
Managerial Framework: Within organizational hierarchy, made possible by grants
Type of Digital Entity: Issues of the online magazine, *HorizonZero*/ZeroHorizon

CS05: Archives of Ontario Web Exhibits
Final Report: Yes Focus: Government
Type of Discipline: Archival
Type of Creator: Provincial archives
Type of Organization: Government body (Province of Ontario)
Managerial Framework: Within governmental hierarchy, under Management Board Secretariat
Type of Digital Entity: Three Web exhibits

CS06: Cybercartographic Atlas of Antarctica
Final Report: Yes Focus: Science
Type of Discipline: Cybercartography
Type of Creator: Geomatics and Cartographic Research Centre (GCRC), Carleton University
Type of Organization: University research group in Ottawa, ON, Canada
Managerial Framework: Granted research group

CS08: Mars Global Surveyor Data Records in the Planetary Data System

Final Report: Yes Focus: Science
Type of Discipline: Space flight
Type of Creator: NASA (National Aeronautics and Space Administration)
Type of Organization: Government agency, headquarters located in Washington, DC, USA
Managerial Framework: Within governmental hierarchy
Type of Digital Entity: Records and data from the Mars Global Surveyor Mission and the Planetary Data System records

CS09(01): Digital Moving Images—Altair4 di Roma

Final Report: Yes Focus: Arts
Type of Discipline: Moving images
Type of Creator: An independent producer (Altair 4)
Type of Organization: Small, private corporation
Managerial Framework: Small, private corporation, run by three partners in Roma, Italy; based on contract
Type of Digital Entity: Multimedia virtual reconstruction of the House of Polybius

CS09(02): Digital Moving Images—National Film Board of Canada

Final Report: Yes Focus: Arts
Type of Discipline: Moving images
Type of Creator: A public filmmaker (National Film Board)
Type of Organization: Government body
Managerial Framework: Within government hierarchy, under the Canadian Heritage Department
Type of Digital Entity: Digital animation products and documentation relating to production

CS09(03): Digital Moving Images—Commercial Film Studio

Final Report: Yes Focus: Arts
Type of Discipline: Moving images
Type of Creator: Anonymous, commercial film studio
Type of Organization: Large, private corporation
Managerial Framework: Large, private film studio
Type of Digital Entity: Artwork related to animated film production

CS09(04): Digital Moving Images—WGBH Boston

Final Report: Yes Focus: Arts
Type of Discipline: Moving images
Type of Creator: A public broadcaster (WGBH)
Type of Organization: Large, public corporation
Managerial Framework: Large, public corporation; in Boston, MA, USA
Type of Digital Entity: Original footage and footage logs generated during the production process of a documentary film

CS10: *The Danube Exodus*: Interactive Multimedia Piece

Final Report: Yes Focus: Arts
Type of Discipline: Multimedia exhibit
Type of Creator: Private individual based in Budapest (installation in Los Angeles, CA, USA)
Type of Organization: Individual, working with an art collective and a research institute
Managerial Framework: Temporary, based on contract or partnership
Type of Digital Entity: Complex media installation

CS12: Antarctic Treaty Searchable Database

Final Report: Yes Focus: Government
Type of Discipline: Treaty documentation
Type of Creator: Private corporation
Type of Organization: Small, private corporation
Managerial Framework: Two individuals, in Ohio, USA
Type of Digital Entity: Database. Purpose: support teaching of Antarctic Treaty documents, enable those searching for Antarctic Treaty materials

CS13: *Obsessed Again...*

Final Report: Yes Focus: Arts
Type of Discipline: Musical performance
Type of Creator: Contract between composer and artist in Vancouver, BC, Canada
Type of Organization: Partnership between composer and artist based on contract
Managerial Framework: Based on contract
Type of Digital Entity: Digital music score

CS14: Archaeological Records in a Geographical Information System: Research in the American Southwest

Final Report: Yes Focus: Science
Type of Discipline: Archaeology
Type of Creator: Center for Desert Archaeology (CDA)
Type of Organization: Private, not-for-profit center located in Tucson, AZ, USA
Managerial Framework: Small, private organization
Type of Digital Entity: GIS Database. Purpose: provide answers to archaeological research questions relating to the aggregation and migration of prehistoric peoples in the American Southwest

CS15: *Waking Dream*

Final Report: Yes Focus: Arts
Type of Discipline: Multimedia performance art
Type of Creator: HCT Laboratory (UBC); three-individuals partnership (digital entity)
Type of Organization: Within university hierarchy
Managerial Framework: Based on partnership
Type of Digital Entity: Web site and multimedia performance art piece

CS17: New York State Department of Motor Vehicles On-line Services System

Final Report: Yes Focus: Government
Type of Discipline: Motor vehicle licensing and driver registration
Type of Creator: New York State Department of Motor Vehicles
Type of Organization: Government department (state)
Managerial Framework: Within governmental hierarchy
Type of Digital Entity: Web site. Purpose: to provide online access to critical state services

CS18: Computerization of Alsace-Moselle's Land Registry

Final Report: Yes Focus: Government
Type of Discipline: Real estate law
Type of Creator: Le Livre Foncier d'Alsace-Moselle; GILFAM (digital entity)
Type of Organization: Within hierarchy of justice system
Managerial Framework: Distributed between offices, judges, clerks

Type of Digital Entity: Database. Purpose: to allow the activities currently underway in the paper-based environment, such as issuing ordinances and completing inscriptions to be done in an automated fashion via a central database

CS19: Preservation and Authentication of Electronic Engineering and Manufacturing Records

Final Report: Yes Focus: Science
Type of Discipline: Engineering and manufacturing
Type of Creator: Various US government departments (Research Division of the Electronic records Archives (ERA), San Diego Supercomputer Center (SDSC), element of the U.S. government with responsibilities in the science, engineering, design and manufacture of complex assemblies)
Type of Organization: Government departments (federal); university unit
Managerial Framework: Within governmental hierarchy; within University of California at San Diego
Type of Digital Entity: Digital engineering and manufacturing records; knowledge-enhanced digital object file

CS20: Revenue On-Line Service (ROS)

Final Report: Yes Focus: Government
Type of Discipline: Tax law
Type of Creator: Office of the Revenue Commissioners of Ireland (Revenue)
Type of Organization: Government body (central)
Managerial Framework: Within governmental hierarchy
Type of Digital Entity: Internet-based tax filing system (Web site)

CS21: Electronic Filing System (EFS) of the Supreme Court of Singapore

Final Report: Yes Focus: Government
Type of Discipline: Law
Type of Creator: Supreme Court of Singapore
Type of Organization: Legal body
Managerial Framework: Within hierarchy of justice system
Type of Digital Entity: Electronic civil and criminal law records filing system

CS22: Electronic Café International: Aging Records from Technology-based Artistic Activities

Final Report: No (unverified draft only) Focus: Arts
Type of Discipline: Multimedia (collaboration and co-creation)
Type of Creator: Electronic Café International; individual artists (digital entities)
Type of Organization: Multimedia international network
Managerial Framework: Two principals, network of artists in Los Angeles, CA, USA
Type of Digital Entity: Accumulation of multimedia related to telecollaborative work

CS24: City of Vancouver Geographic Information System (VanMap)

Final Report: Yes Focus: Government
Type of Discipline:
Type of Creator: City of Vancouver, BC, Canada
Type of Organization: Municipal government
Managerial Framework: Team, within government hierarchy (IT Department)
Type of Digital Entity: GIS database. Purpose: to allow the City of Vancouver to “meet the needs of internal users in providing services to Vancouver’s citizens and businesses”

CS25: Legacoop of Bologna Web Site

Final Report: Yes Focus: Government

Type of Discipline: Cooperative

Type of Creator: Cooperative network

Type of Organization: Provincial body of cooperative network in Bologna, Italy

Managerial Framework: Divided into departments, within network hierarchy

Type of Digital Entity: Web site. Purpose: to increase communication with and maintain the cooperative network of Legacoop Bologna's members

CS26: Microvariability & Oscillations of Stars (MOST) Satellite Mission - Preservation of Space Telescope Data

Final Report: Yes Focus: Science

Type of Discipline: Astronomy

Type of Creator: Microvariablity and Oscillations of STars satellite mission

Type of Organization: Partnership between Canada Space Agency, industry, universities

Managerial Framework: Based on partnership

Type of Digital Entity: Space telescope data and engineering telemetry

APPENDIX 12

Domain 2 Research Questions

Domain 2 Research Questions

- What does record reliability mean in the context of artistic, scientific and governmental activities? To what extent can the electronic records created in the course of each type of activity be considered reliable and why? What requirements on their form and controls on their creation would make us presume that they are reliable?

- What does record accuracy mean in the context of each activity? To what extent can the electronic records created in the course of each type of activity be considered accurate and why? What controls on their creation would make us presume that these records are accurate?

- What does authenticity mean in the context of each activity? To what extent is the definition of record authenticity adopted by InterPARES 1 relevant to the records resulting from each type of activity and from the use of increasingly complex digital technology?

- On what basis can the records created in the course of each activity be presumed authentic? How, in the absence of such presumption, can their authenticity be verified?

- How is the authenticity of these records affected by their transmission across space and time? What controls on the process of transmission would ensure that these records will continue to be recognized as authentic?

- Are the conceptual requirements for reliability and authenticity developed by the UBC-MAS project and InterPARES 1 for administrative and legal records generated within databases and document management systems applicable to the records studied by InterPARES 2?

- Do the participants in electronic transactions have shared access to reliable and accurate information about the terms and effects of the transactions? What would constitute reliable and accurate records of transactions in current electronic service delivery initiatives?

- What would be the consequence of issuing guidelines for record creation on the nature of the records of each activity?

- How can cultural differences, freedom of expression, freedom of inquiry, and right to privacy be reflected in those guidelines?

- What technological and intellectual tools would assist creators to generate records that can be authentically preserved over time?

- What legal or moral obligations exist regarding the creation, use and preservation of the records under investigation?

APPENDIX 13

Domain 3 Template for Case Study Analysis

Domain 3 Template for Case Study Analysis

Domain 3
Case Study Analysis

Case Study
Title
Organization

1. What types of entities does the diplomatic analysis identify in this case study? (i.e. records, publications, data, etc.)[1]
   1a. If there are no records, should there be records? If not, why not?
   1b. If there should be records, what kinds of records should be created to satisfy the creator's needs (as defined by an archivist)?
   1c. What characteristics of records (as defined by an archivist)[2] are missing yet necessary to preserve these entities?
      - completed as part of an action
      - involving a communication among 3 juridical or physical persons (e.g. author, writer, addressee), or over time
      - a fixed documentary form
      - a stable content
      - an archival bond with other records either inside or outside the system
      - an identifiable context
2. Are the entities reliable? If not, why not? (Give evidence from both the diplomatic analysis and the case study report.)
3. Are the entities accurate? If not, why not? (Give evidence from both the diplomatic analysis and the case study report.)
4. To what degree can the entities be presumed to be authentic, and why?[3] (The answer to this question requires providing the evidence for all benchmark requirements that have been fulfilled and also reaching a cumulative presumption of authenticity. The higher the number of satisfied requirements, and the greater the degree to which requirement is satisfied, the stronger the presumption of authenticity.)

   Benchmark Requirements Supporting the Production of Authentic Copies of Electronic Records (these apply to the creator):
      1. Capture of identity and integrity metadata
      2. Enforcement of access privileges
      3. Protection against loss and corruption

[1] If multiple entity types are identified, answer questions 1a onward for each type of entity selected for analysis. See Appendix [1a], section 1 for a discussion of General vs. Special Diplomatics and section 2 for a number of definitions relevant to Question 1.
[2] See Appendix [1a], section 3 for a more complete definition of the characteristics of a record.
[3] See Appendix [1b] for a more extensive discussion of the elements of the Benchmark and Baseline Requirements Supporting the Presumption of Authenticity of Electronic Records, from InterPARES 1.

4. Protection against media and technology obsolescence
5. Established documentary forms
6. Ability to authenticate records
7. Procedures in place to identify the authoritative record
8. Procedures in place to properly document removal and transfer of records from the creator's originating system

Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records (these apply to the preserver):

1. Controls over records transfer, maintenance, and reproduction
2. Documentation of reproduction process and its effects
3. Archival description

5. For what purpose(s) are the entities to be preserved?[4]
6. Has the feasibility of preservation been explored?
   6a. If yes, what elements and components need to be preserved?
7. Which preservation strategies[5] might most usefully be applied, and what are their strengths and weaknesses, including costs and degree of technical difficulty?
   7a. Which alternative preservation strategies[6] might be applied? What are their strengths and weaknesses, including costs and degree of technical difficulty?
8. What additional information does the preserver need to know to facilitate appraisal and preservation?
   8a. If required information is missing, where should it come from and how should it be made manifest?
9. Are there any policies in place that affect preservation?
   9a. Are there any policies in place that present obstacles to preservation?
   9b. Are there any policies that would need to be put in place to facilitate appraisal and preservation?

---

[4] If multiple purposes are identified, answer questions #3 onwards for each purpose.

[5] For a list of the many existing preservation strategies, see the *Domain 3 Preservation Strategies* document. Draft 3, dated 2006-02-22 is located at ip2(d3)_preservation_strategies.pdf. Confirm that you have the most up-to-date version of the document. Identify any of the maintenance strategies from Section A that are not undertaken or are prevented by the existing recordkeeping system. Then apply any one or combination of more than one strategy from Section B that might be applied by the preserver.

[6] For a list of the many existing preservation strategies, see the *Domain 3 Preservation Strategies* document. Draft 3, dated 2006-02-22 is located at ip2(d3)_preservation_strategies.pdf. Confirm that you have the most up-to-date version of the document. Identify any of the maintenance strategies from Section A that are not undertaken or are prevented by the existing recordkeeping system. Then apply any one or combination of more than one strategy from Section B that might be applied by the preserver.

Appendix [Definitions]

1. General vs. Special Diplomatics

The limitations of the diplomatic model of a record as it is elaborated in the *Template for Analysis* are attributable mainly to the fact that the model was built on the premises of general diplomatics. *General diplomatics* seeks to decontextualize records, to eliminate their particularities, variations and anomalies in the interest of identifying the common, shared elements of records that cut across juridical, provenancial, and technological boundaries. Given the complexity and variety of electronic systems, it might make more sense to adopt and adapt the approach of *special diplomatics*, which, traditionally, has focused on the records of individual chanceries and specific juridical systems. In such an approach, one would begin with an analysis of the various features of the systems themselves and the broader record-keeping environment in their own terms, with all their particularities, variations, and anomalies; and, on the basis of that analysis, begin to build a more general framework.[7]

2. Current IP2 Definitions

Record[8]
[Archival Science] - *n.,* A document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference.

Reliable record[9]
[Archival Science] - *n.,* A record capable of standing for the facts to which it attests.

Reliability refers to the trustworthiness of a record as a statement of fact. It exists when a record can stand for the fact it is about, and is established by examining the completeness of the record's form and the amount of control exercised on the process of its creation. The records forms generated using new information technologies make increasingly difficult to determine when a record is complete and whether the controls established on its creation are either sufficient or effective for anyone to be able to assume its reliability.[10]

Accuracy refers to the truthfulness of the content of the record and can only be established through content analysis. With administrative and legal records, it is usually inferred on the basis of the degree of the records' reliability and is only verified when such degree is very low. The volatility of the digital medium, the ease of change, editing, and the difficulty of version control, all make it harder to presume accuracy on the traditional [basis].[11]

[7] From the Final Report of the Authenticity Task Force, p. 24, at http://www.interpares.org/display_file.cfm?doc=ip1_atf_report.pdf
[8] From the InterPARES Glossary, at http://www.interpares.org/ip2/ip2_terminology_db.cfm.
[9] Ibid.
[10] From the detailed proposal, p. 12, at http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf.
[11] Ibid.

Authenticity refers to the trustworthiness of a record as a record. An authentic record is one that is what it purports to be and has not been tampered with or otherwise corrupted. Authenticity is established by assessing the identity and the integrity of the record.[12]

3. Characteristics of a Record:[13]

- a **fixed documentary form**, which means that:

  - the binary content of the record, including indicators of its documentary form, are stored in a manner that ensures it remains complete and unaltered

  - technology has been maintained and procedures defined and enforced to ensure that the content is presented or rendered with the same documentary form it had when it was set aside

- a **stable content**

- an **archival bond** with other records either inside or outside the system

- an **identifiable context**, which means that it participates in or supports an action, either procedurally or as part of the decision-making process (meaning its creation may be mandatory or discretionary), and at least three persons (author, writer, and addressee) are involved in its creation (although these three conceptual persons may in fact be only one physical or juridical person).

Appendix [Benchmark and Baseline Requirements]

[Note: This section has been omitted. Abridged versions of the InterPARES 1 Benchmark and Baseline Requirements are instead provided in Appendices 22a and 22b, respectively.

[12] Ibid.
[13] From the Final Report of the Authenticity Task Force, InterPARES 1, p. 6, at http://www.interpares.org/display_file.cfm?doc=ip1_atf_report.pdf.

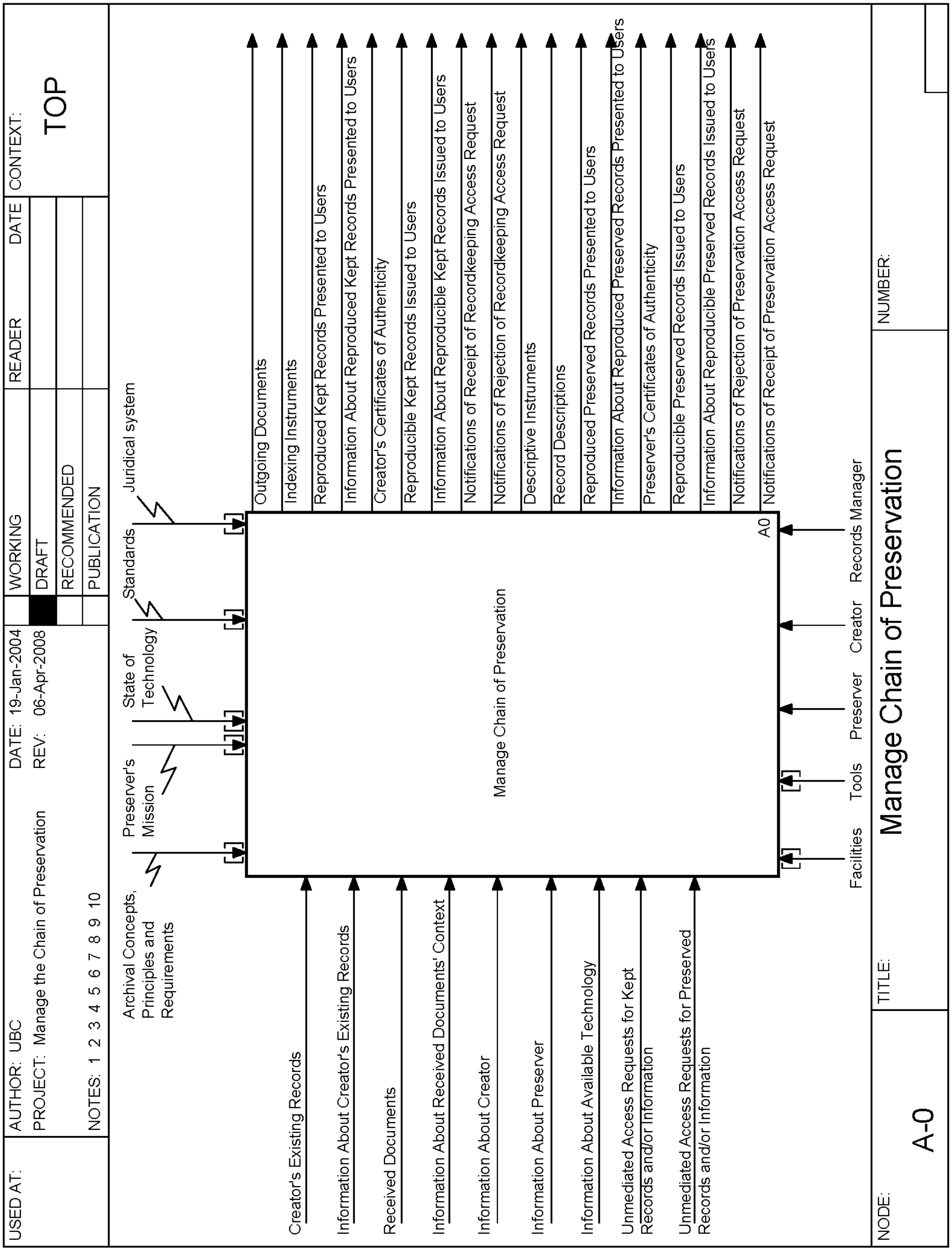
APPENDIX 14

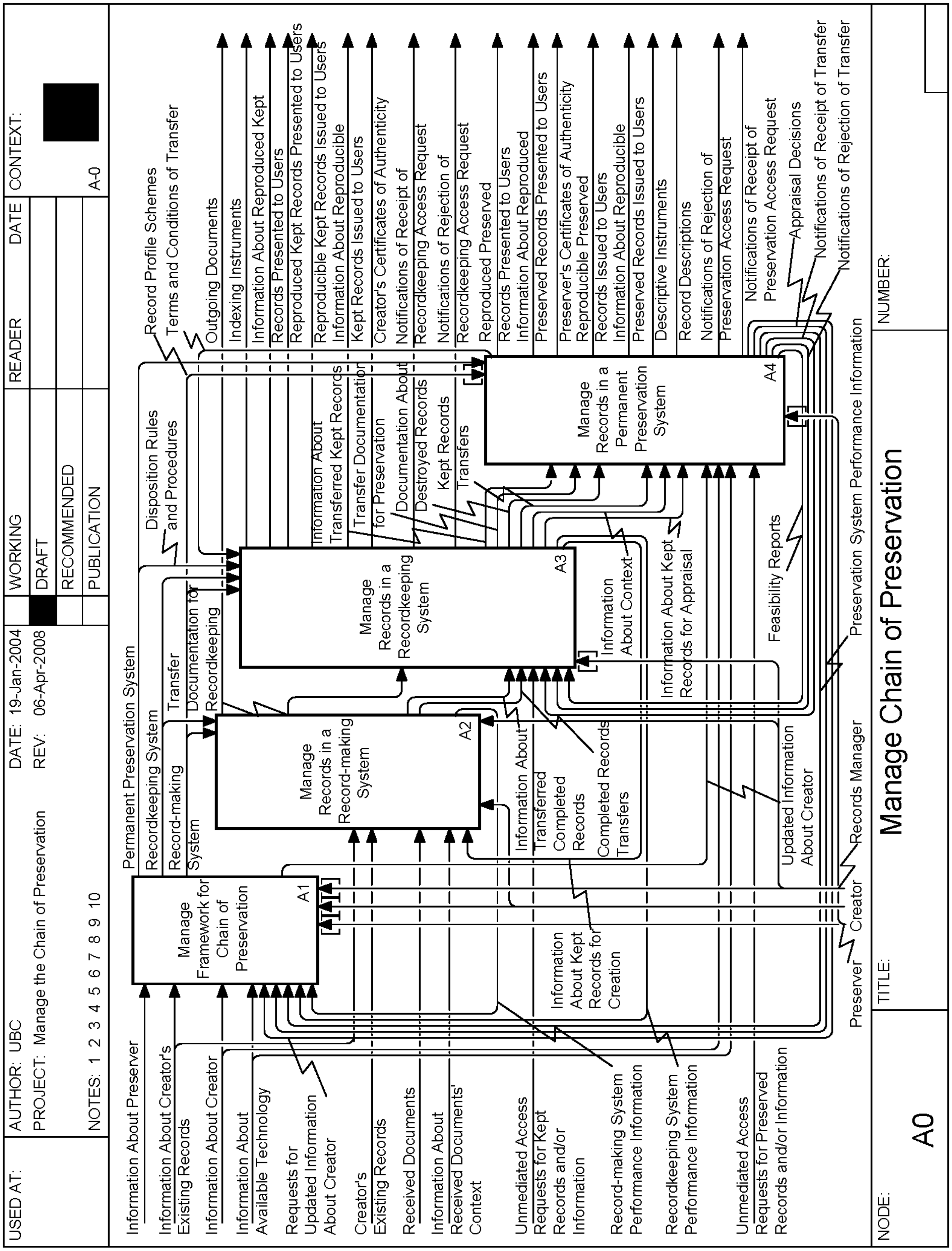
Chain of Preservation Model

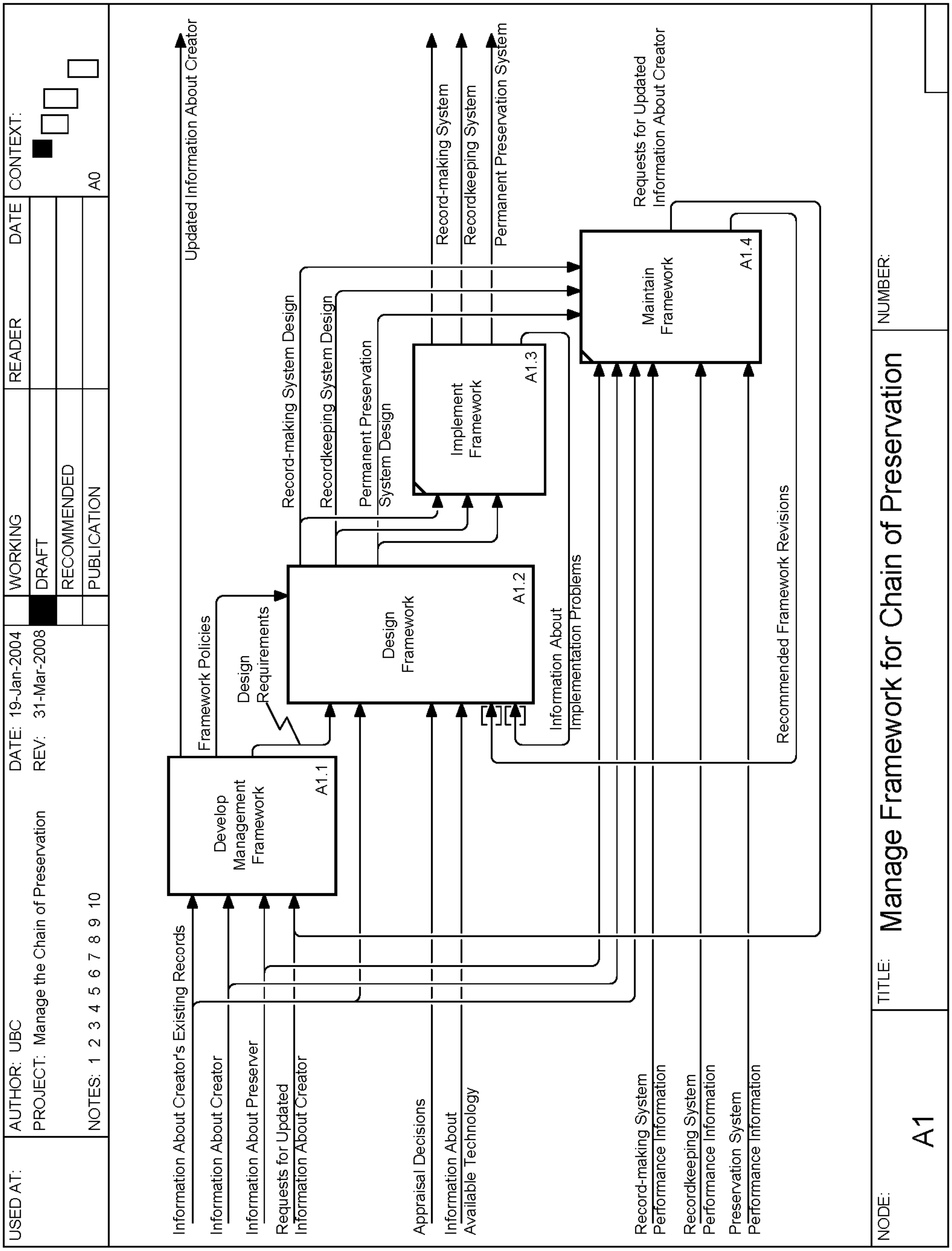
Diagrams and Definitions

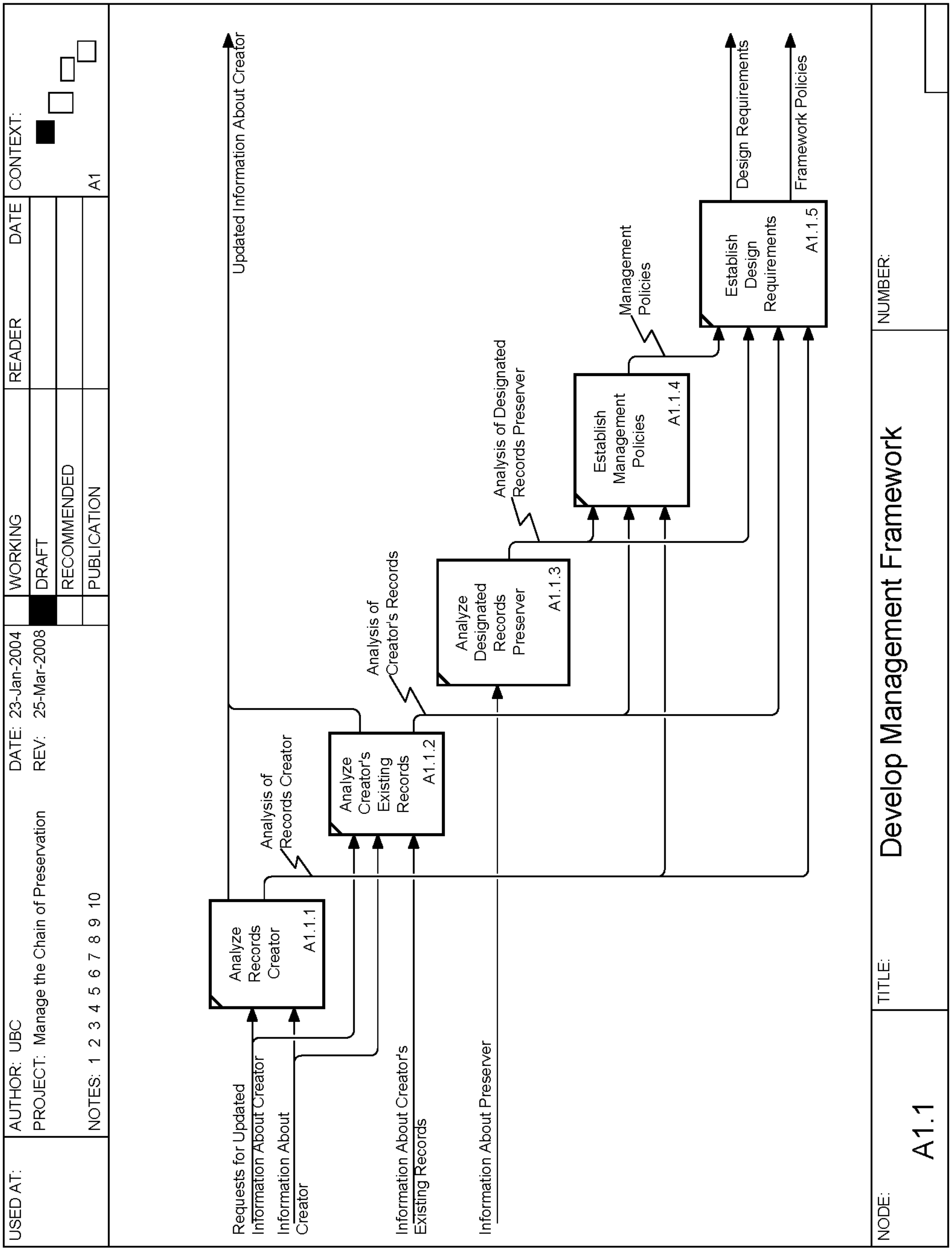
by

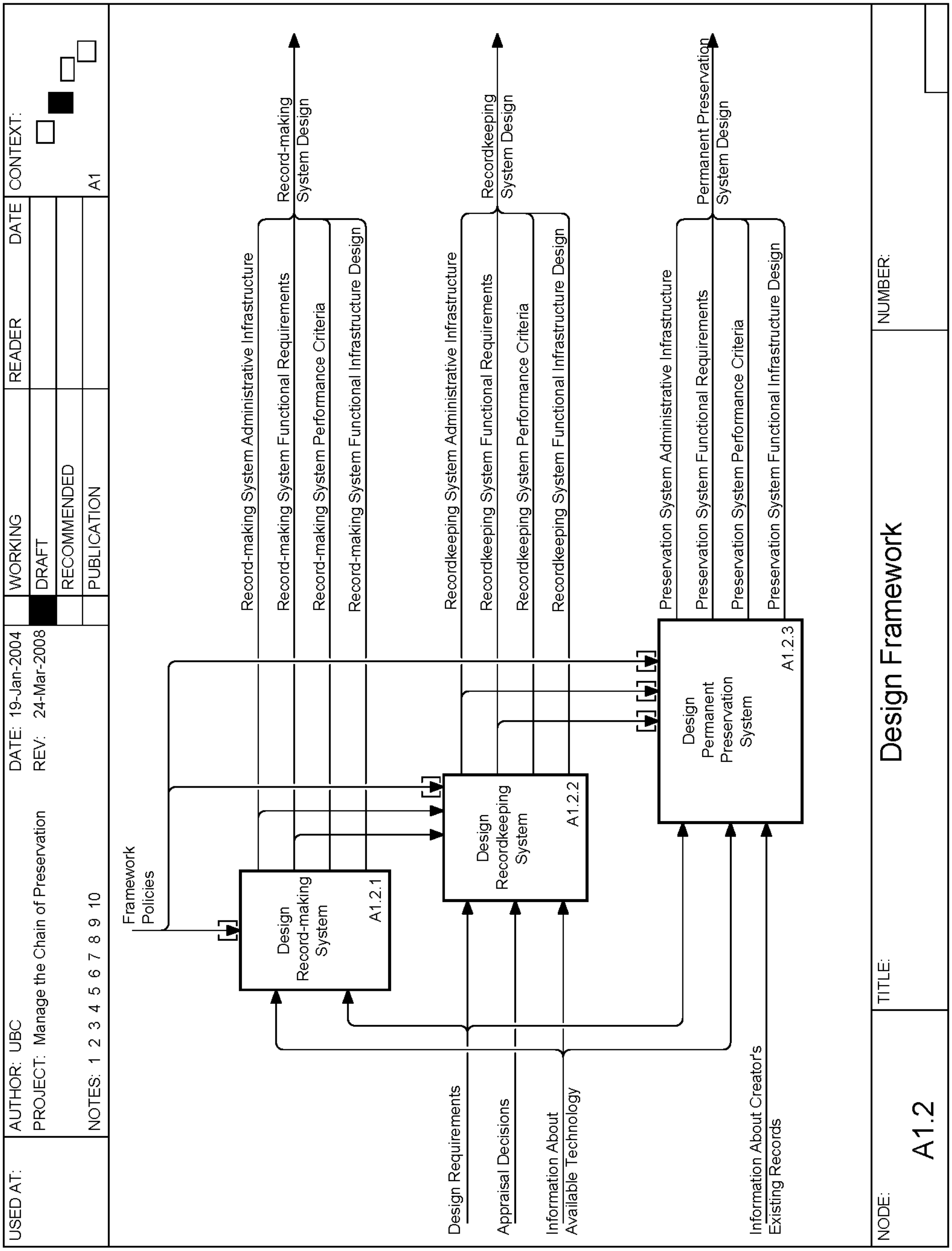
Terry Eastwood, Bart Ballaux, Rachel Mills and Randy Preston
School of Library, Archival and Information Studies,
The University of British Columbia

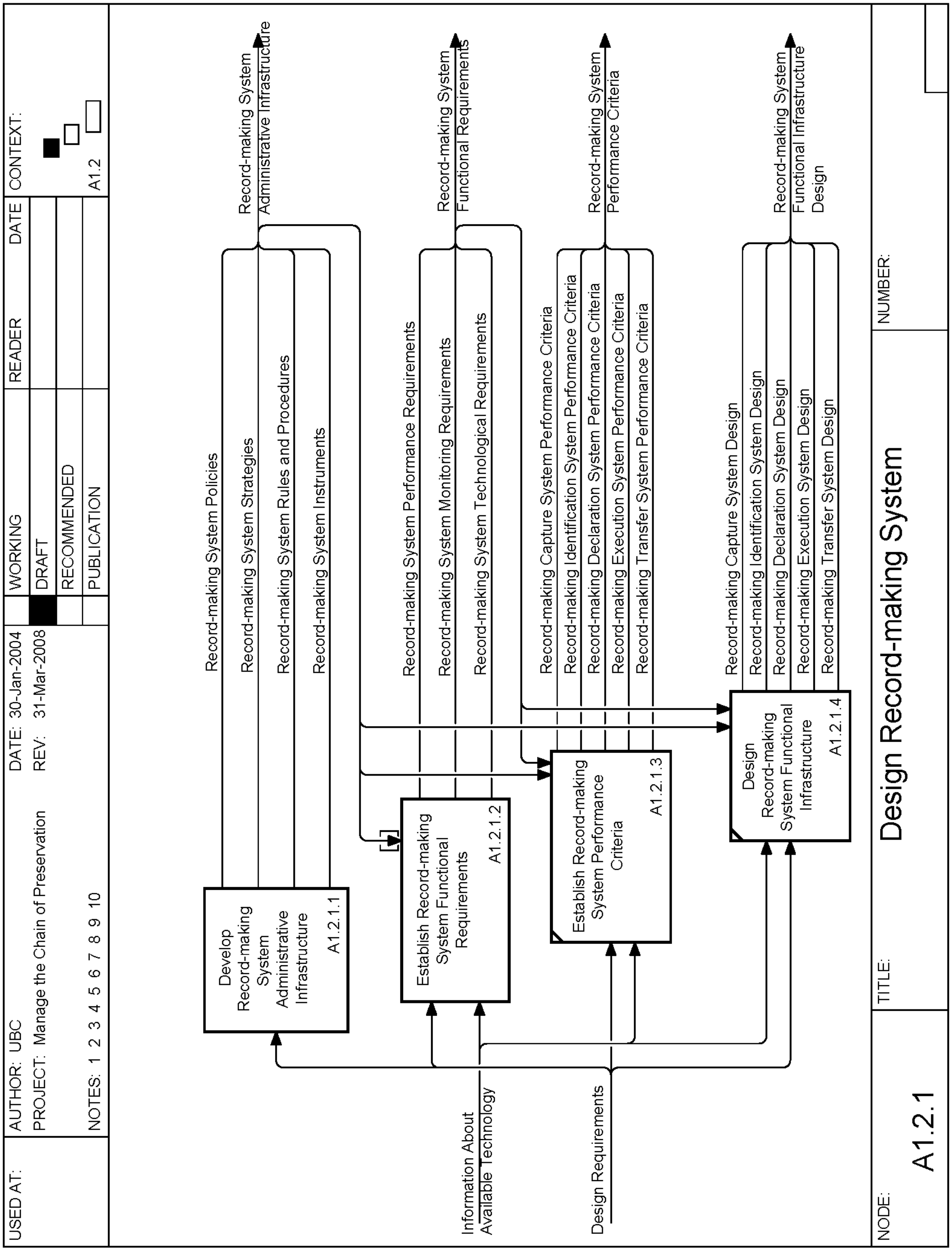

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 19-Jan-2004
REV: 06-Apr-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
TOP
Archival Concepts, Principles and Requirements
Preserver's Mission
State of Technology
Standards
Juridical system
Creator's Existing Records
Information About Creator's Existing Records
Received Documents
Information About Received Documents' Context
Information About Creator
Information About Preserver
Information About Available Technology
Unmediated Access Requests for Kept Records and/or Information
Unmediated Access Requests for Preserved Records and/or Information
Manage Chain of Preservation
A0
Outgoing Documents
Indexing Instruments
Reproduced Kept Records Presented to Users
Information About Reproduced Kept Records Presented to Users
Creator's Certificates of Authenticity
Reproducible Kept Records Issued to Users
Information About Reproducible Kept Records Issued to Users
Notifications of Receipt of Recordkeeping Access Request
Notifications of Rejection of Recordkeeping Access Request
Descriptive Instruments
Record Descriptions
Reproduced Preserved Records Presented to Users
Information About Reproduced Preserved Records Presented to Users
Preserver's Certificates of Authenticity
Reproducible Preserved Records Issued to Users
Information About Reproducible Preserved Records Issued to Users
Notifications of Rejection of Preservation Access Request
Notifications of Receipt of Preservation Access Request
Facilities
Tools
Preserver
Creator
Records Manager
NODE:
A-0
TITLE:
Manage Chain of Preservation
NUMBER:

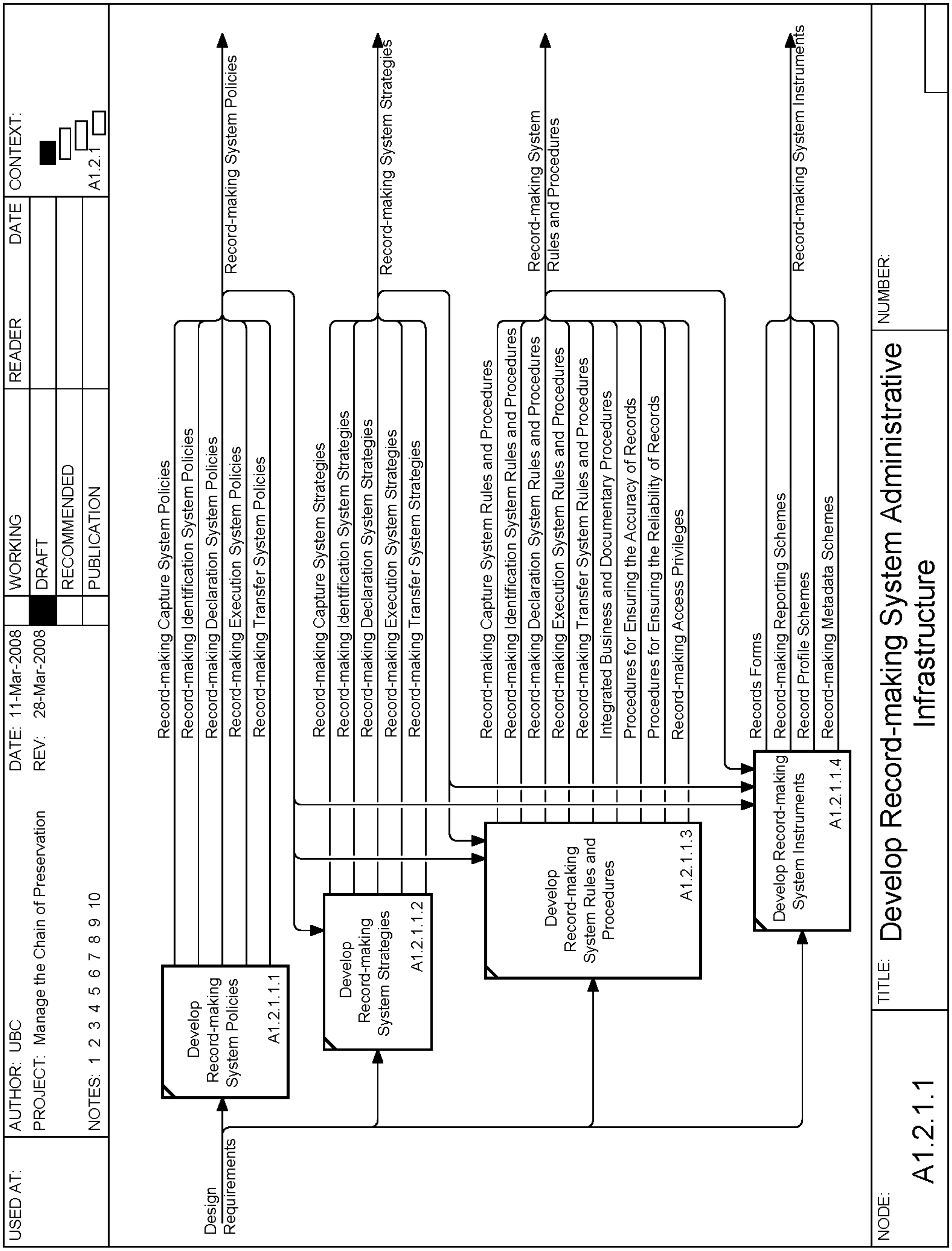

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 19-Jan-2004
REV: 06-Apr-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A-0
Information About Preserver
Information About Creator's Existing Records
Information About Creator
Information About Available Technology
Requests for Updated Information About Creator
Creator's Existing Records
Received Documents
Information About Received Documents' Context
Unmediated Access Requests for Kept Records and/or Information
Record-making System Performance Information
Recordkeeping System Performance Information
Unmediated Access Requests for Preserved Records and/or Information
Manage Framework for Chain of Preservation
A1
Manage Records in a Record-making System
A2
Manage Records in a Recordkeeping System
A3
Manage Records in a Permanent Preservation System
A4
Permanent Preservation System
Recordkeeping System
Record-making System
Transfer
Documentation for Recordkeeping
Disposition Rules and Procedures
Record Profile Schemes
Terms and Conditions of Transfer
Information About Kept Records for Creation
Information About Transferred Completed Records
Completed Records Transfers
Information About Context
Information About Kept Records for Appraisal
Feasibility Reports
Updated Information About Creator
Information About Transferred Kept Records
Transfer Documentation for Preservation
Documentation About Destroyed Records
Kept Records Transfers
Preservation System Performance Information
Preserver
Creator
Records Manager
Outgoing Documents
Indexing Instruments
Information About Reproduced Kept Records Presented to Users
Reproduced Kept Records Presented to Users
Reproducible Kept Records Issued to Users
Information About Reproducible Kept Records Issued to Users
Creator's Certificates of Authenticity
Notifications of Receipt of Recordkeeping Access Request
Notifications of Rejection of Recordkeeping Access Request
Reproduced Preserved Records Presented to Users
Information About Reproduced Preserved Records Presented to Users
Preserver's Certificates of Authenticity
Reproducible Preserved Records Issued to Users
Information About Reproducible Preserved Records Issued to Users
Descriptive Instruments
Record Descriptions
Notifications of Rejection of Preservation Access Request
Notifications of Receipt of Preservation Access Request
Appraisal Decisions
Notifications of Receipt of Transfer
Notifications of Rejection of Transfer
NODE:
A0
TITLE:
Manage Chain of Preservation
NUMBER:

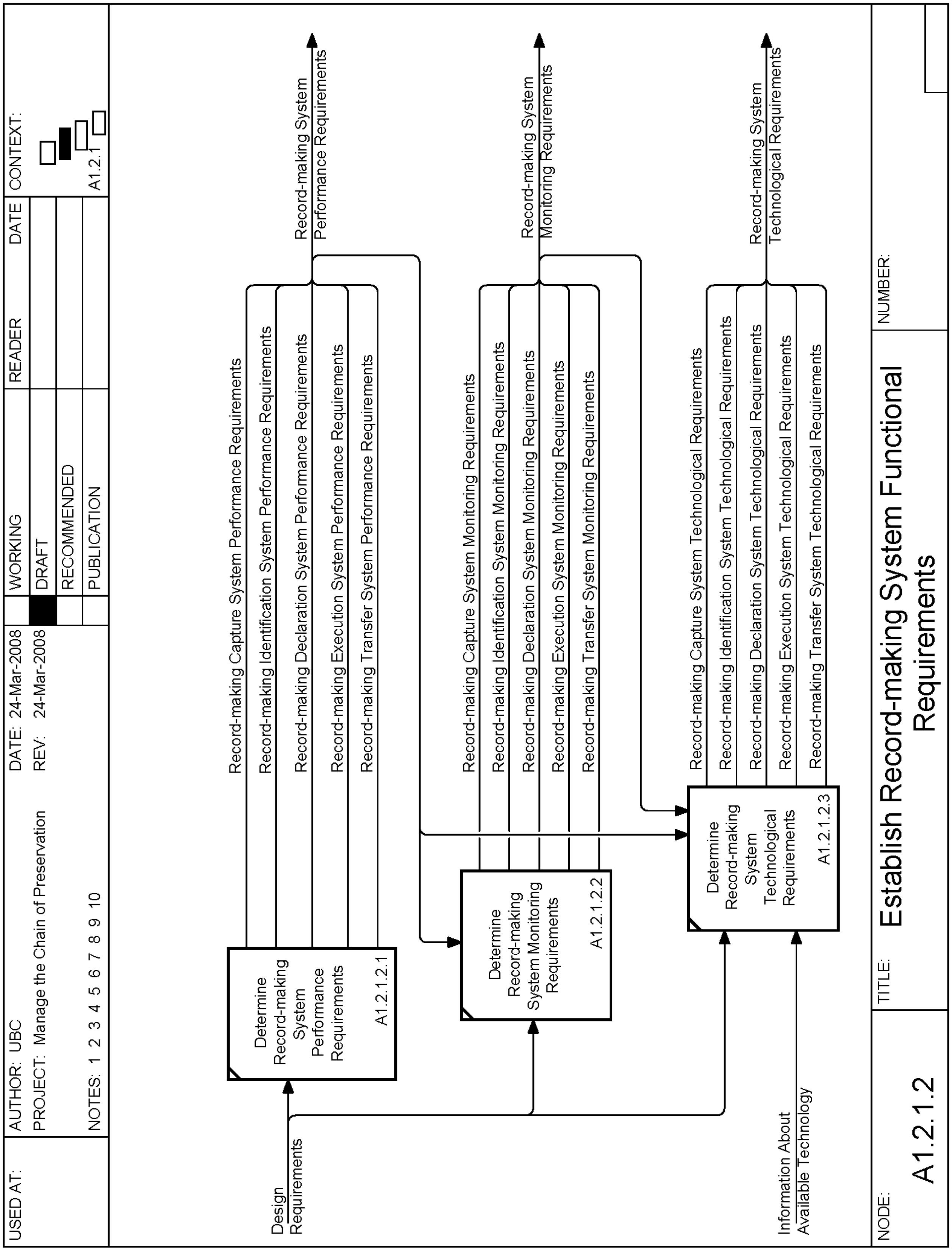

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 19-Jan-2004
REV: 31-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A0
Information About Creator's Existing Records
Information About Creator
Information About Preserver
Requests for Updated Information About Creator
Appraisal Decisions
Information About Available Technology
Record-making System Performance Information
Recordkeeping System Performance Information
Preservation System Performance Information
Develop Management Framework
A1.1
Framework Policies
Design Requirements
Design Framework
A1.2
Information About Implementation Problems
Recommended Framework Revisions
Record-making System Design
Recordkeeping System Design
Permanent Preservation System Design
Implement Framework
A1.3
Maintain Framework
A1.4
Updated Information About Creator
Record-making System
Recordkeeping System
Permanent Preservation System
Requests for Updated Information About Creator
NODE:
A1
TITLE:
Manage Framework for Chain of Preservation
NUMBER:

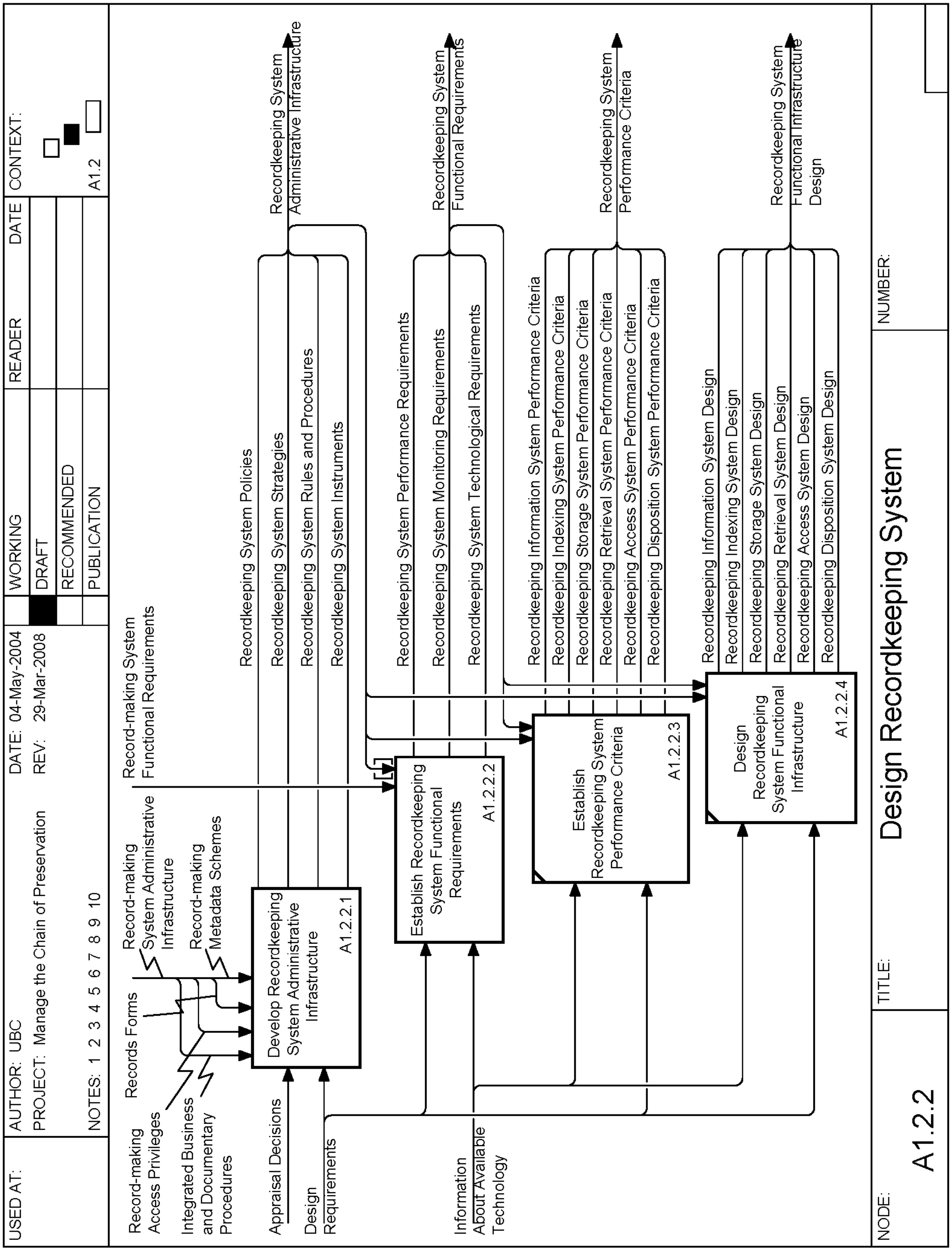

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 23-Jan-2004
REV: 25-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1
Requests for Updated Information About Creator
Information About Creator
Information About Creator's Existing Records
Information About Preserver
Analyze Records Creator
A1.1.1
Analysis of Records Creator
Analyze Creator's Existing Records
A1.1.2
Updated Information About Creator
Analysis of Creator's Records
Analyze Designated Records Preserver
A1.1.3
Analysis of Designated Records Preserver
Establish Management Policies
A1.1.4
Management Policies
Establish Design Requirements
A1.1.5
Design Requirements
Framework Policies
NODE:
A1.1
TITLE:
Develop Management Framework
NUMBER:

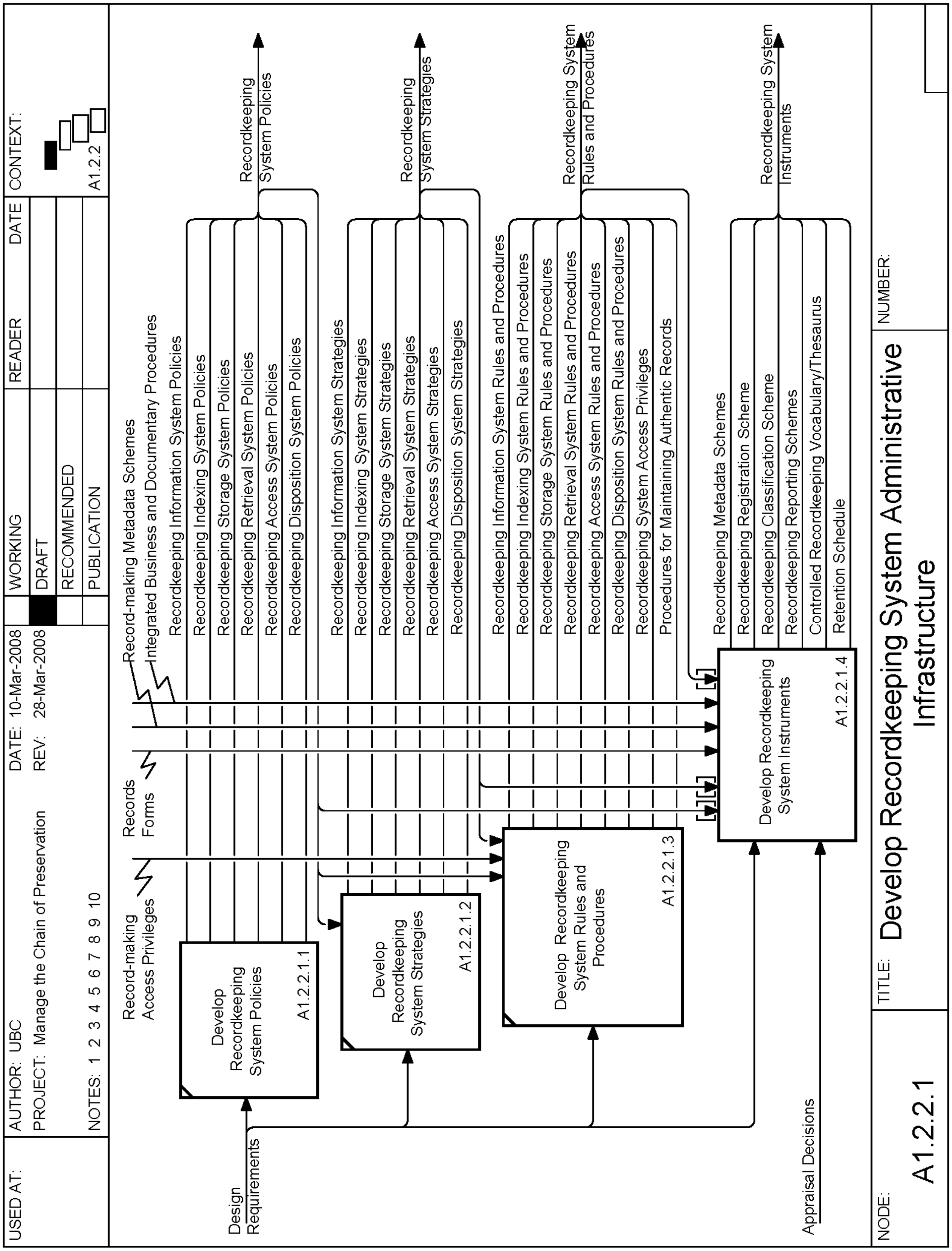

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 19-Jan-2004
REV: 24-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1
Framework Policies
Design Record-making System
A1.2.1
Design Recordkeeping System
A1.2.2
Design Permanent Preservation System
A1.2.3
Record-making System Administrative Infrastructure
Record-making System Functional Requirements
Record-making System Performance Criteria
Record-making System Functional Infrastructure Design
Record-making System Design
Recordkeeping System Administrative Infrastructure
Recordkeeping System Functional Requirements
Recordkeeping System Performance Criteria
Recordkeeping System Functional Infrastructure Design
Recordkeeping System Design
Preservation System Administrative Infrastructure
Preservation System Functional Requirements
Preservation System Performance Criteria
Preservation System Functional Infrastructure Design
Permanent Preservation System Design
Design Requirements
Appraisal Decisions
Information About Available Technology
Information About Creator's Existing Records
NODE:
A1.2
TITLE:
Design Framework
NUMBER:

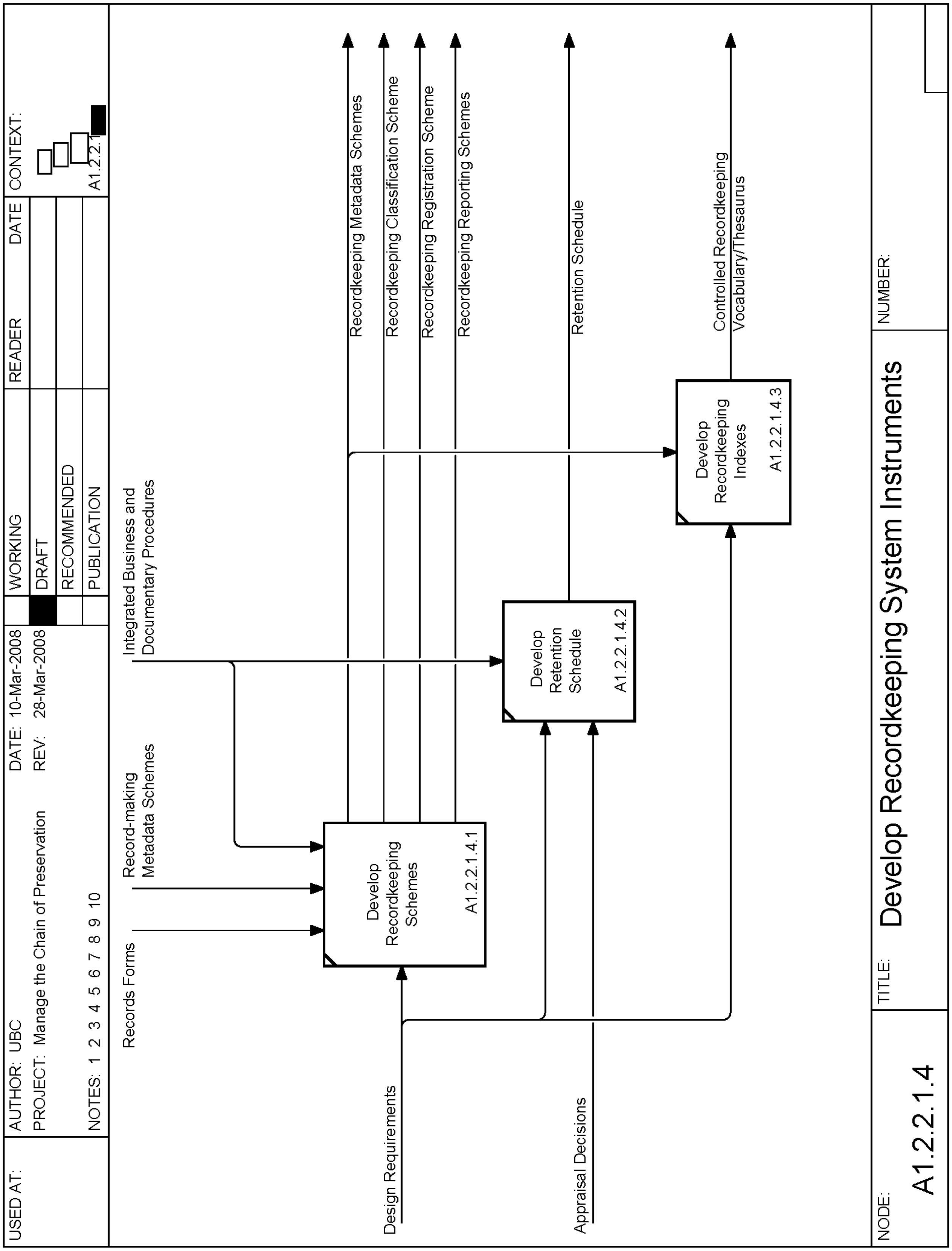

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 30-Jan-2004
REV: 31-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.2
Information About Available Technology
Design Requirements
Develop Record-making System Administrative Infrastructure
A1.2.1.1
Establish Record-making System Functional Requirements
A1.2.1.2
Establish Record-making System Performance Criteria
A1.2.1.3
Design Record-making System Functional Infrastructure
A1.2.1.4
Record-making System Policies
Record-making System Strategies
Record-making System Rules and Procedures
Record-making System Instruments
Record-making System Performance Requirements
Record-making System Monitoring Requirements
Record-making System Technological Requirements
Record-making Capture System Performance Criteria
Record-making Identification System Performance Criteria
Record-making Declaration System Performance Criteria
Record-making Execution System Performance Criteria
Record-making Transfer System Performance Criteria
Record-making Capture System Design
Record-making Identification System Design
Record-making Declaration System Design
Record-making Execution System Design
Record-making Transfer System Design
Record-making System Administrative Infrastructure
Record-making System Functional Requirements
Record-making System Performance Criteria
Record-making System Functional Infrastructure Design
NODE:
A1.2.1
TITLE:
Design Record-making System
NUMBER:

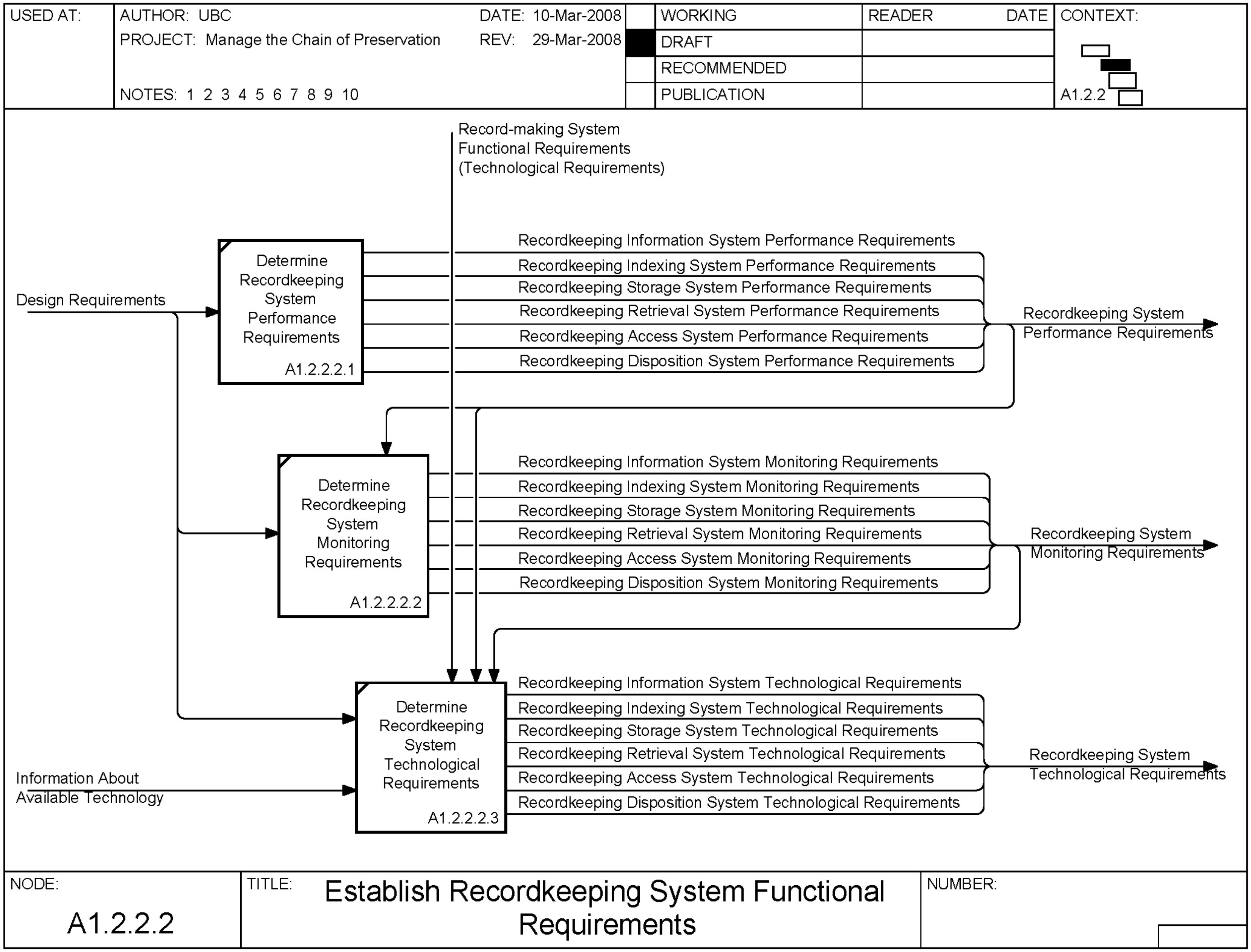

USED AT:
AUTHOR: UBC
DATE: 11-Mar-2008
REV: 28-Mar-2008
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.2.1
Design Requirements
Develop Record-making System Policies
A1.2.1.1.1
Develop Record-making System Strategies
A1.2.1.1.2
Develop Record-making System Rules and Procedures
A1.2.1.1.3
Develop Record-making System Instruments
A1.2.1.1.4
Record-making Capture System Policies
Record-making Identification System Policies
Record-making Declaration System Policies
Record-making Execution System Policies
Record-making Transfer System Policies
Record-making System Policies
Record-making Capture System Strategies
Record-making Identification System Strategies
Record-making Declaration System Strategies
Record-making Execution System Strategies
Record-making Transfer System Strategies
Record-making System Strategies
Record-making Capture System Rules and Procedures
Record-making Identification System Rules and Procedures
Record-making Declaration System Rules and Procedures
Record-making Execution System Rules and Procedures
Record-making Transfer System Rules and Procedures
Integrated Business and Documentary Procedures
Procedures for Ensuring the Accuracy of Records
Procedures for Ensuring the Reliability of Records
Record-making Access Privileges
Record-making System Rules and Procedures
Records Forms
Record-making Reporting Schemes
Record Profile Schemes
Record-making Metadata Schemes
Record-making System Instruments
NODE:
A1.2.1.1
TITLE:
Develop Record-making System Administrative Infrastructure
NUMBER:


USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 24-Mar-2008
REV: 24-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.2.1
Design Requirements
Information About Available Technology
Determine Record-making System Performance Requirements
A1.2.1.2.1
Determine Record-making System Monitoring Requirements
A1.2.1.2.2
Determine Record-making System Technological Requirements
A1.2.1.2.3
Record-making Capture System Performance Requirements
Record-making Identification System Performance Requirements
Record-making Declaration System Performance Requirements
Record-making Execution System Performance Requirements
Record-making Transfer System Performance Requirements
Record-making System Performance Requirements
Record-making Capture System Monitoring Requirements
Record-making Identification System Monitoring Requirements
Record-making Declaration System Monitoring Requirements
Record-making Execution System Monitoring Requirements
Record-making Transfer System Monitoring Requirements
Record-making System Monitoring Requirements
Record-making Capture System Technological Requirements
Record-making Identification System Technological Requirements
Record-making Declaration System Technological Requirements
Record-making Execution System Technological Requirements
Record-making Transfer System Technological Requirements
Record-making System Technological Requirements
NODE:
A1.2.1.2
TITLE:
Establish Record-making System Functional Requirements
NUMBER:

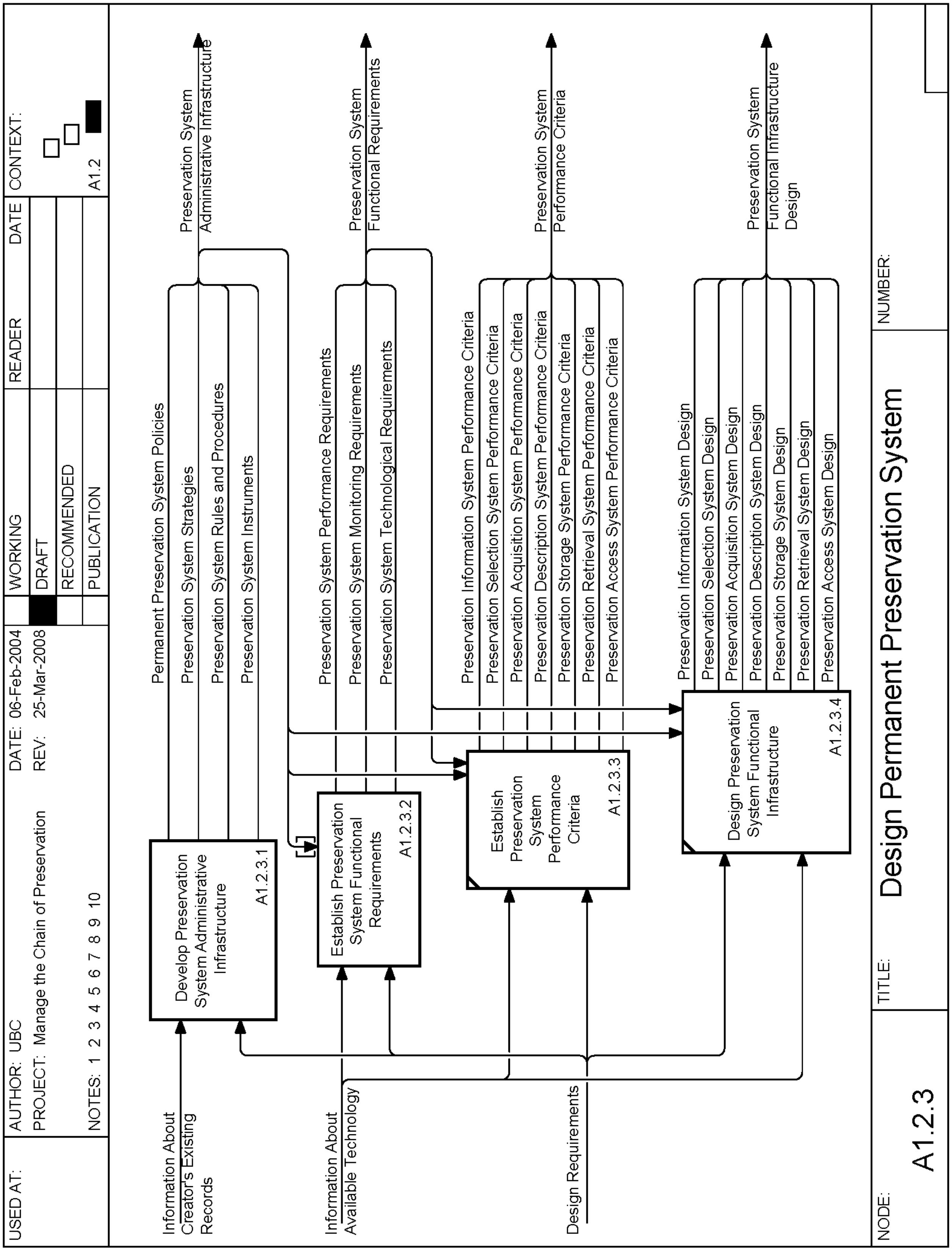

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 04-May-2004
REV: 29-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.2
Record-making Access Privileges
Integrated Business and Documentary Procedures
Appraisal Decisions
Design Requirements
Information About Available Technology
Records Forms
Record-making System Administrative Infrastructure
Record-making Metadata Schemes
Record-making System Functional Requirements
Develop Recordkeeping System Administrative Infrastructure
A1.2.2.1
Establish Recordkeeping System Functional Requirements
A1.2.2.2
Establish Recordkeeping System Performance Criteria
A1.2.2.3
Design Recordkeeping System Functional Infrastructure
A1.2.2.4
Recordkeeping System Policies
Recordkeeping System Strategies
Recordkeeping System Rules and Procedures
Recordkeeping System Instruments
Recordkeeping System Performance Requirements
Recordkeeping System Monitoring Requirements
Recordkeeping System Technological Requirements
Recordkeeping Information System Performance Criteria
Recordkeeping Indexing System Performance Criteria
Recordkeeping Storage System Performance Criteria
Recordkeeping Retrieval System Performance Criteria
Recordkeeping Access System Performance Criteria
Recordkeeping Disposition System Performance Criteria
Recordkeeping Information System Design
Recordkeeping Indexing System Design
Recordkeeping Storage System Design
Recordkeeping Retrieval System Design
Recordkeeping Access System Design
Recordkeeping Disposition System Design
Recordkeeping System Administrative Infrastructure
Recordkeeping System Functional Requirements
Recordkeeping System Performance Criteria
Recordkeeping System Functional Infrastructure Design
NODE:
A1.2.2
TITLE:
Design Recordkeeping System
NUMBER:

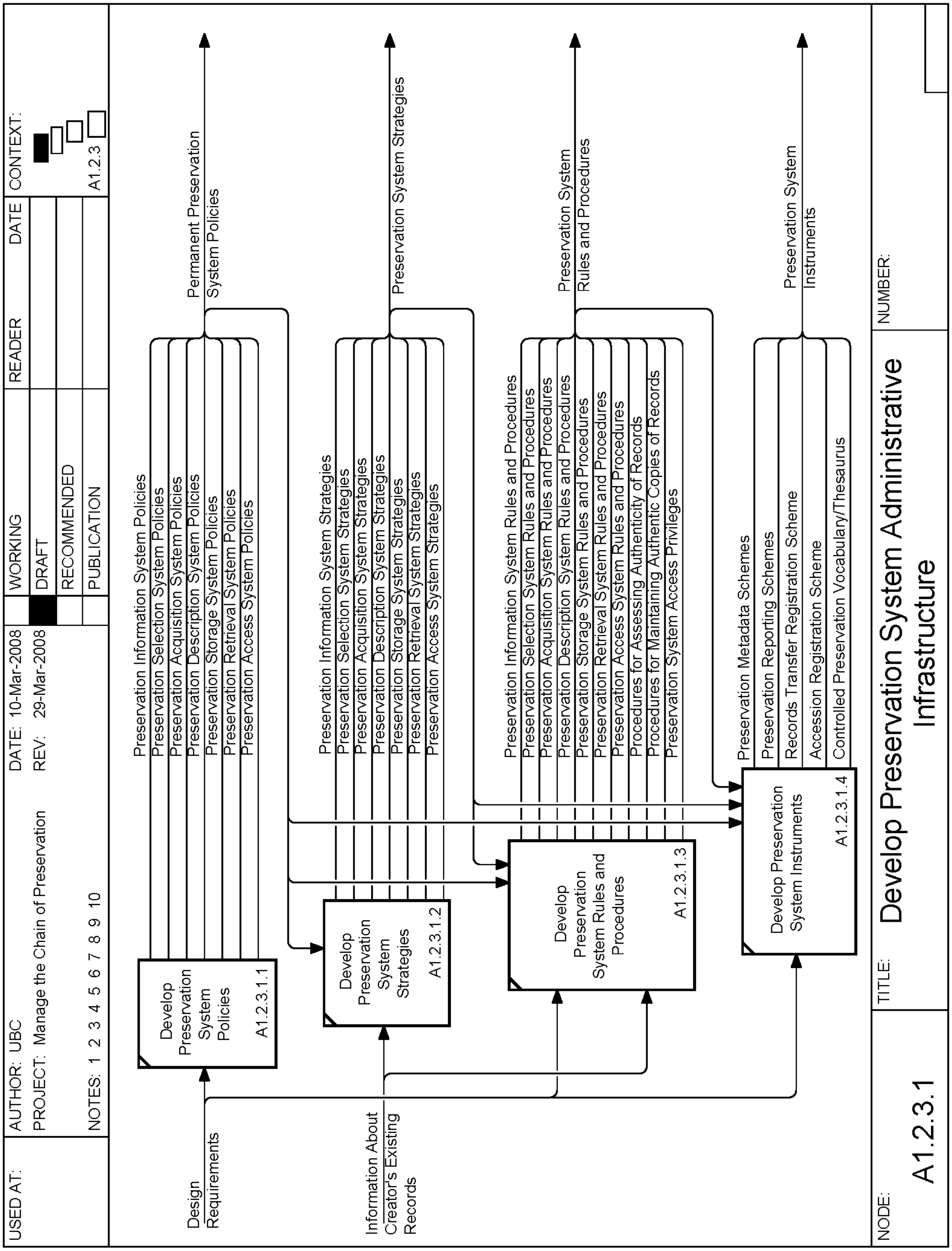

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 10-Mar-2008
REV: 28-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.2.2
Design Requirements
Record-making Access Privileges
Records Forms
Record-making Metadata Schemes
Integrated Business and Documentary Procedures
Recordkeeping Information System Policies
Recordkeeping Indexing System Policies
Recordkeeping Storage System Policies
Recordkeeping Retrieval System Policies
Recordkeeping Access System Policies
Recordkeeping Disposition System Policies
Recordkeeping Information System Strategies
Recordkeeping Indexing System Strategies
Recordkeeping Storage System Strategies
Recordkeeping Retrieval System Strategies
Recordkeeping Access System Strategies
Recordkeeping Disposition System Strategies
Recordkeeping Information System Rules and Procedures
Recordkeeping Indexing System Rules and Procedures
Recordkeeping Storage System Rules and Procedures
Recordkeeping Retrieval System Rules and Procedures
Recordkeeping Access System Rules and Procedures
Recordkeeping Disposition System Rules and Procedures
Recordkeeping System Access Privileges
Procedures for Maintaining Authentic Records
Recordkeeping Metadata Schemes
Recordkeeping Registration Scheme
Recordkeeping Classification Scheme
Recordkeeping Reporting Schemes
Controlled Recordkeeping Vocabulary/Thesaurus
Retention Schedule
Recordkeeping System Policies
Recordkeeping System Strategies
Recordkeeping System Rules and Procedures
Recordkeeping System Instruments
Develop Recordkeeping System Policies
A1.2.2.1.1
Develop Recordkeeping System Strategies
A1.2.2.1.2
Develop Recordkeeping System Rules and Procedures
A1.2.2.1.3
Develop Recordkeeping System Instruments
A1.2.2.1.4
Appraisal Decisions
NODE:
A1.2.2.1
TITLE:
Develop Recordkeeping System Administrative Infrastructure
NUMBER:


USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 10-Mar-2008
REV: 28-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.2.2.1
Records Forms
Record-making Metadata Schemes
Integrated Business and Documentary Procedures
Develop Recordkeeping Schemes
A1.2.2.1.4.1
Develop Retention Schedule
A1.2.2.1.4.2
Develop Recordkeeping Indexes
A1.2.2.1.4.3
Recordkeeping Metadata Schemes
Recordkeeping Classification Scheme
Recordkeeping Registration Scheme
Recordkeeping Reporting Schemes
Retention Schedule
Controlled Recordkeeping Vocabulary/Thesaurus
Design Requirements
Appraisal Decisions
NODE:
A1.2.2.1.4
TITLE:
Develop Recordkeeping System Instruments
NUMBER:

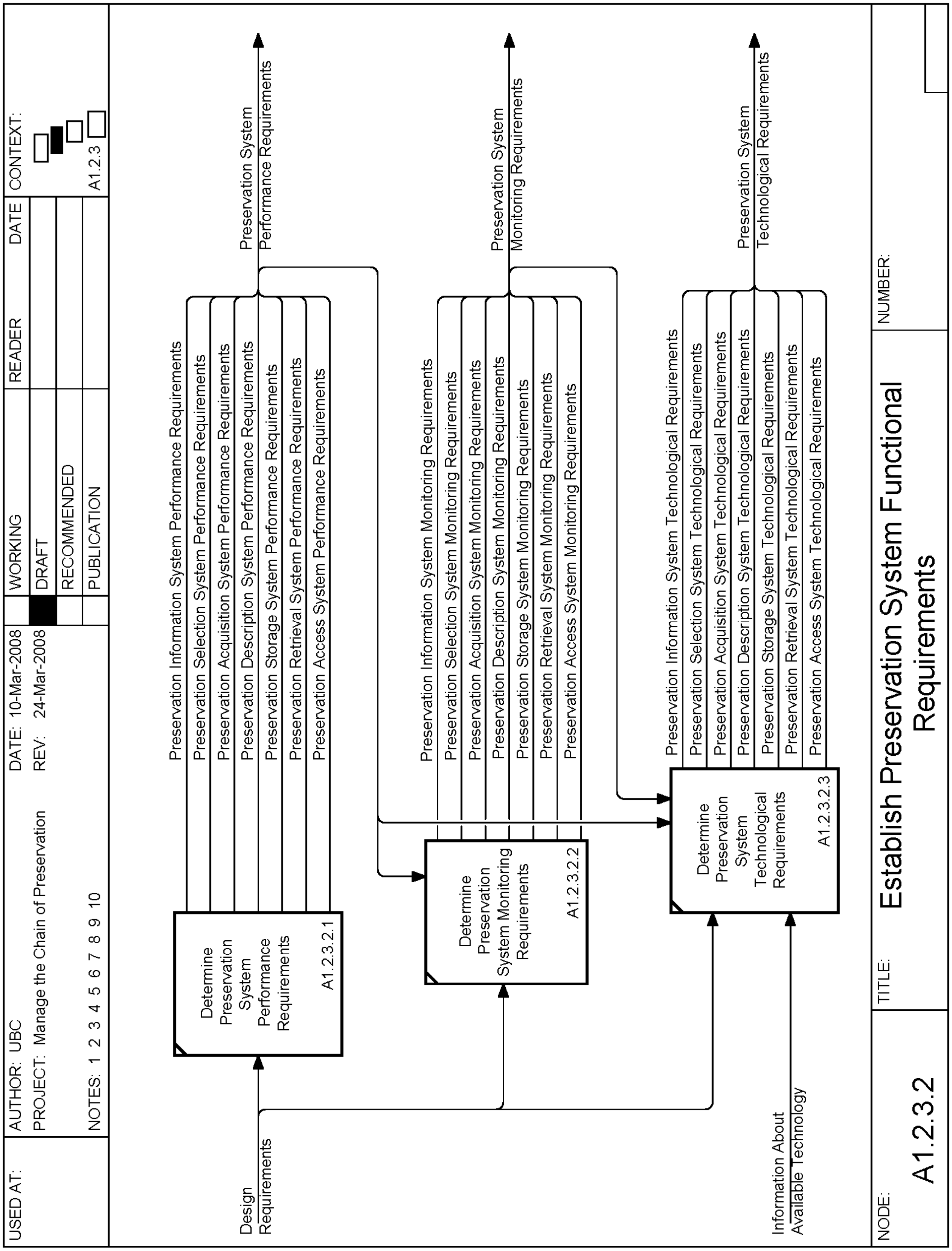


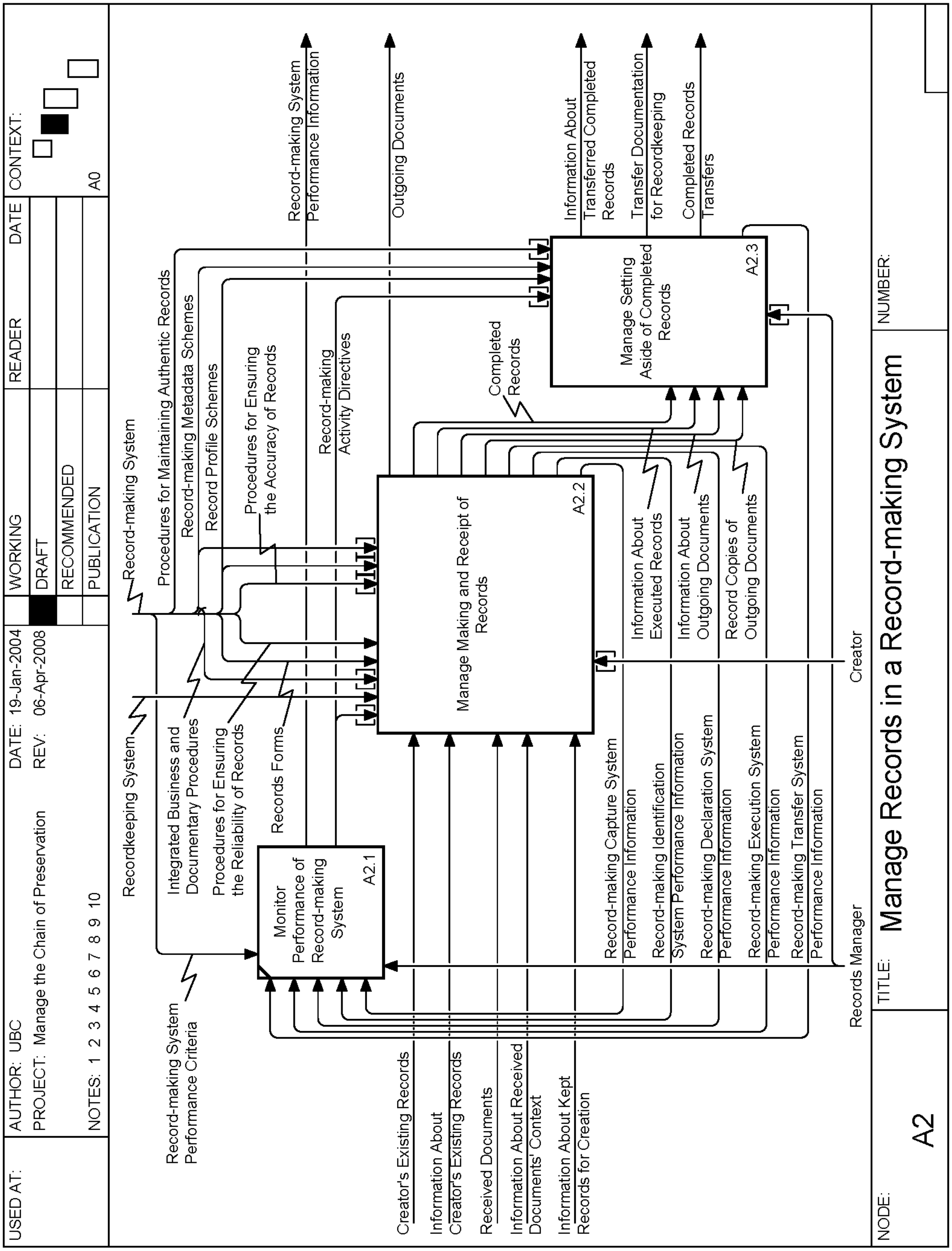
USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
DATE: 10-Mar-2008
REV: 29-Mar-2008
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.2.2
Record-making System Functional Requirements (Technological Requirements)
Design Requirements
Determine Recordkeeping System Performance Requirements
A1.2.2.2.1
Recordkeeping Information System Performance Requirements
Recordkeeping Indexing System Performance Requirements
Recordkeeping Storage System Performance Requirements
Recordkeeping Retrieval System Performance Requirements
Recordkeeping Access System Performance Requirements
Recordkeeping Disposition System Performance Requirements
Recordkeeping System Performance Requirements
Determine Recordkeeping System Monitoring Requirements
A1.2.2.2.2
Recordkeeping Information System Monitoring Requirements
Recordkeeping Indexing System Monitoring Requirements
Recordkeeping Storage System Monitoring Requirements
Recordkeeping Retrieval System Monitoring Requirements
Recordkeeping Access System Monitoring Requirements
Recordkeeping Disposition System Monitoring Requirements
Recordkeeping System Monitoring Requirements
Information About Available Technology
Determine Recordkeeping System Technological Requirements
A1.2.2.2.3
Recordkeeping Information System Technological Requirements
Recordkeeping Indexing System Technological Requirements
Recordkeeping Storage System Technological Requirements
Recordkeeping Retrieval System Technological Requirements
Recordkeeping Access System Technological Requirements
Recordkeeping Disposition System Technological Requirements
Recordkeeping System Technological Requirements
NODE:
A1.2.2.2
TITLE:
Establish Recordkeeping System Functional Requirements
NUMBER:

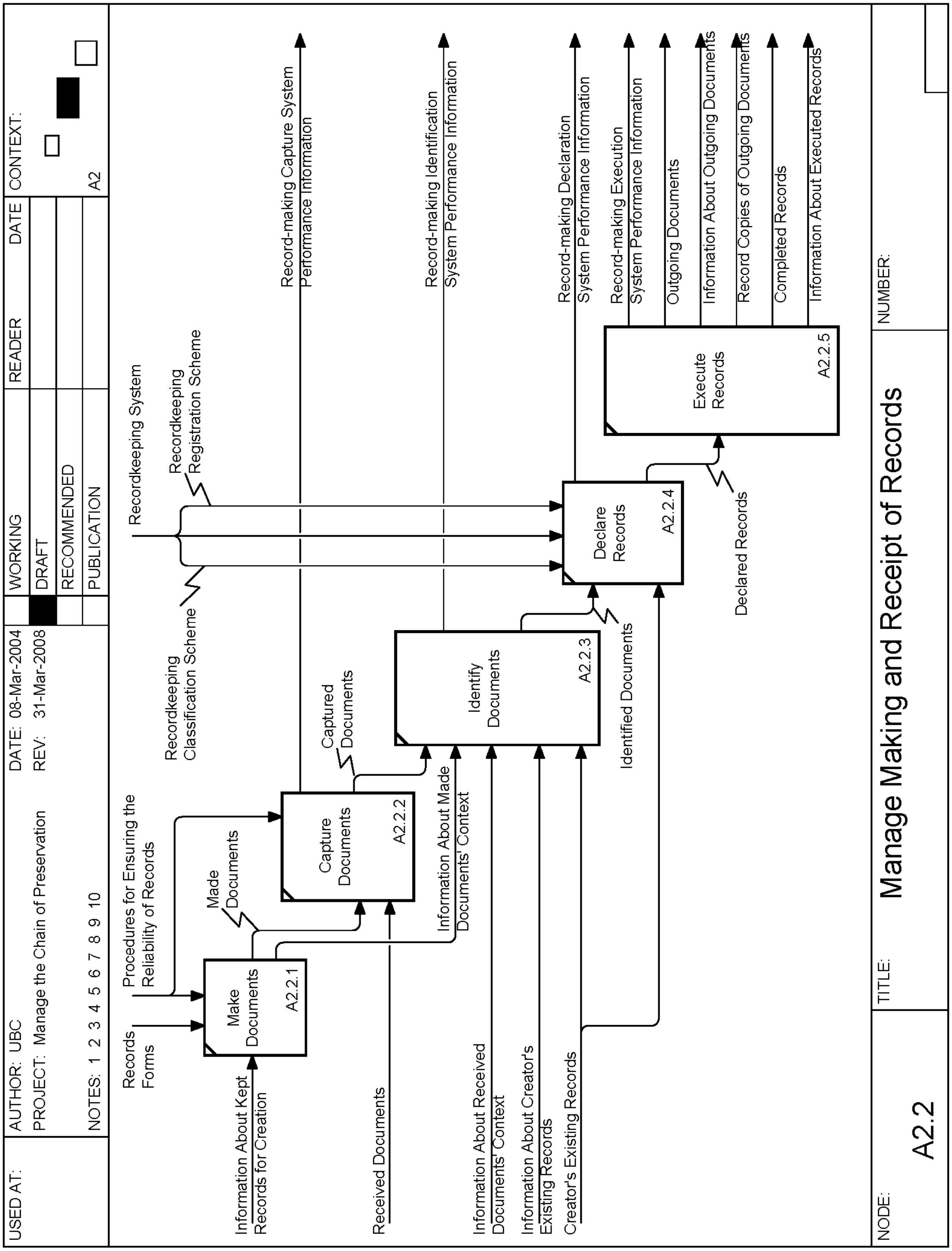

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 06-Feb-2004
REV: 25-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.2
Information About Creator's Existing Records
Information About Available Technology
Design Requirements
Permanent Preservation System Policies
Preservation System Strategies
Preservation System Rules and Procedures
Preservation System Instruments
Preservation System Performance Requirements
Preservation System Monitoring Requirements
Preservation System Technological Requirements
Preservation Information System Performance Criteria
Preservation Selection System Performance Criteria
Preservation Acquisition System Performance Criteria
Preservation Description System Performance Criteria
Preservation Storage System Performance Criteria
Preservation Retrieval System Performance Criteria
Preservation Access System Performance Criteria
Preservation Information System Design
Preservation Selection System Design
Preservation Acquisition System Design
Preservation Description System Design
Preservation Storage System Design
Preservation Retrieval System Design
Preservation Access System Design
Preservation System Administrative Infrastructure
Preservation System Functional Requirements
Preservation System Performance Criteria
Preservation System Functional Infrastructure Design
Develop Preservation System Administrative Infrastructure
A1.2.3.1
Establish Preservation System Functional Requirements
A1.2.3.2
Establish Preservation System Performance Criteria
A1.2.3.3
Design Preservation System Functional Infrastructure
A1.2.3.4
NODE:
A1.2.3
TITLE:
Design Permanent Preservation System
NUMBER:

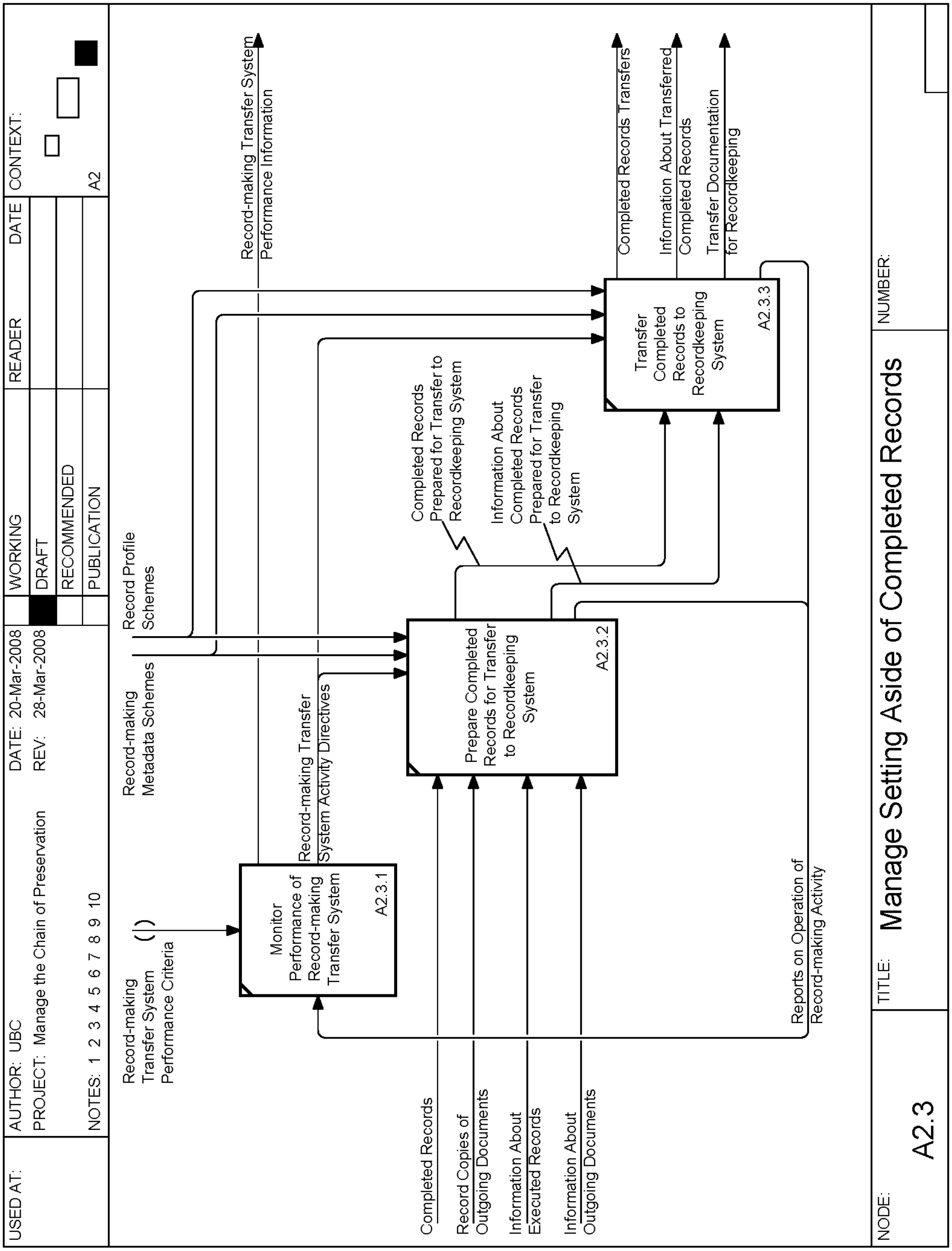


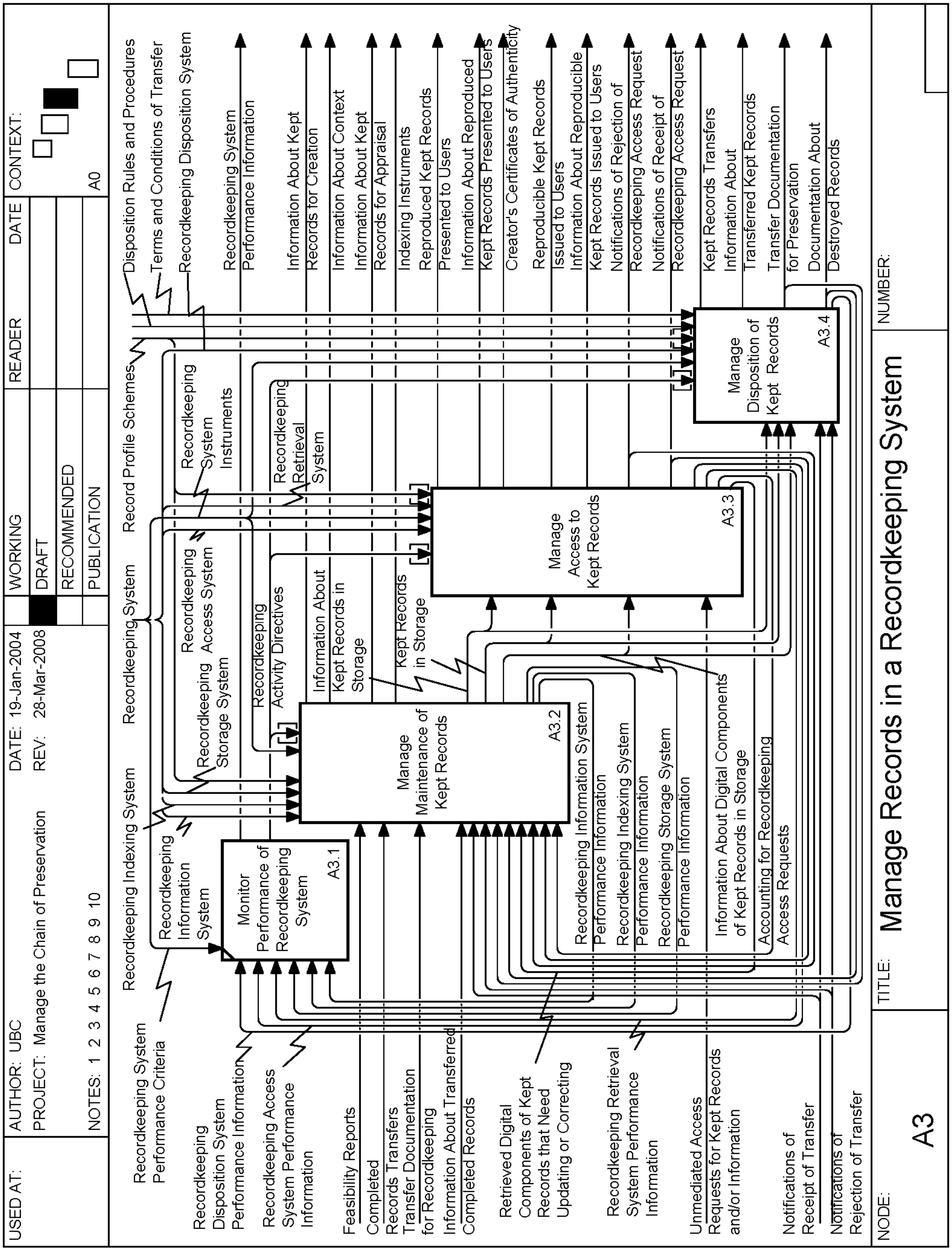
USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 10-Mar-2008
REV: 29-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.2.3
Design Requirements
Information About Creator's Existing Records
Develop Preservation System Policies
A1.2.3.1.1
Develop Preservation System Strategies
A1.2.3.1.2
Develop Preservation System Rules and Procedures
A1.2.3.1.3
Develop Preservation System Instruments
A1.2.3.1.4
Preservation Information System Policies
Preservation Selection System Policies
Preservation Acquisition System Policies
Preservation Description System Policies
Preservation Storage System Policies
Preservation Retrieval System Policies
Preservation Access System Policies
Permanent Preservation System Policies
Preservation Information System Strategies
Preservation Selection System Strategies
Preservation Acquisition System Strategies
Preservation Description System Strategies
Preservation Storage System Strategies
Preservation Retrieval System Strategies
Preservation Access System Strategies
Preservation System Strategies
Preservation Information System Rules and Procedures
Preservation Selection System Rules and Procedures
Preservation Acquisition System Rules and Procedures
Preservation Description System Rules and Procedures
Preservation Storage System Rules and Procedures
Preservation Retrieval System Rules and Procedures
Preservation Access System Rules and Procedures
Procedures for Assessing Authenticity of Records
Procedures for Maintaining Authentic Copies of Records
Preservation System Access Privileges
Preservation System Rules and Procedures
Preservation Metadata Schemes
Preservation Reporting Schemes
Records Transfer Registration Scheme
Accession Registration Scheme
Controlled Preservation Vocabulary/Thesaurus
Preservation System Instruments
NODE:
A1.2.3.1
TITLE:
Develop Preservation System Administrative Infrastructure
NUMBER:

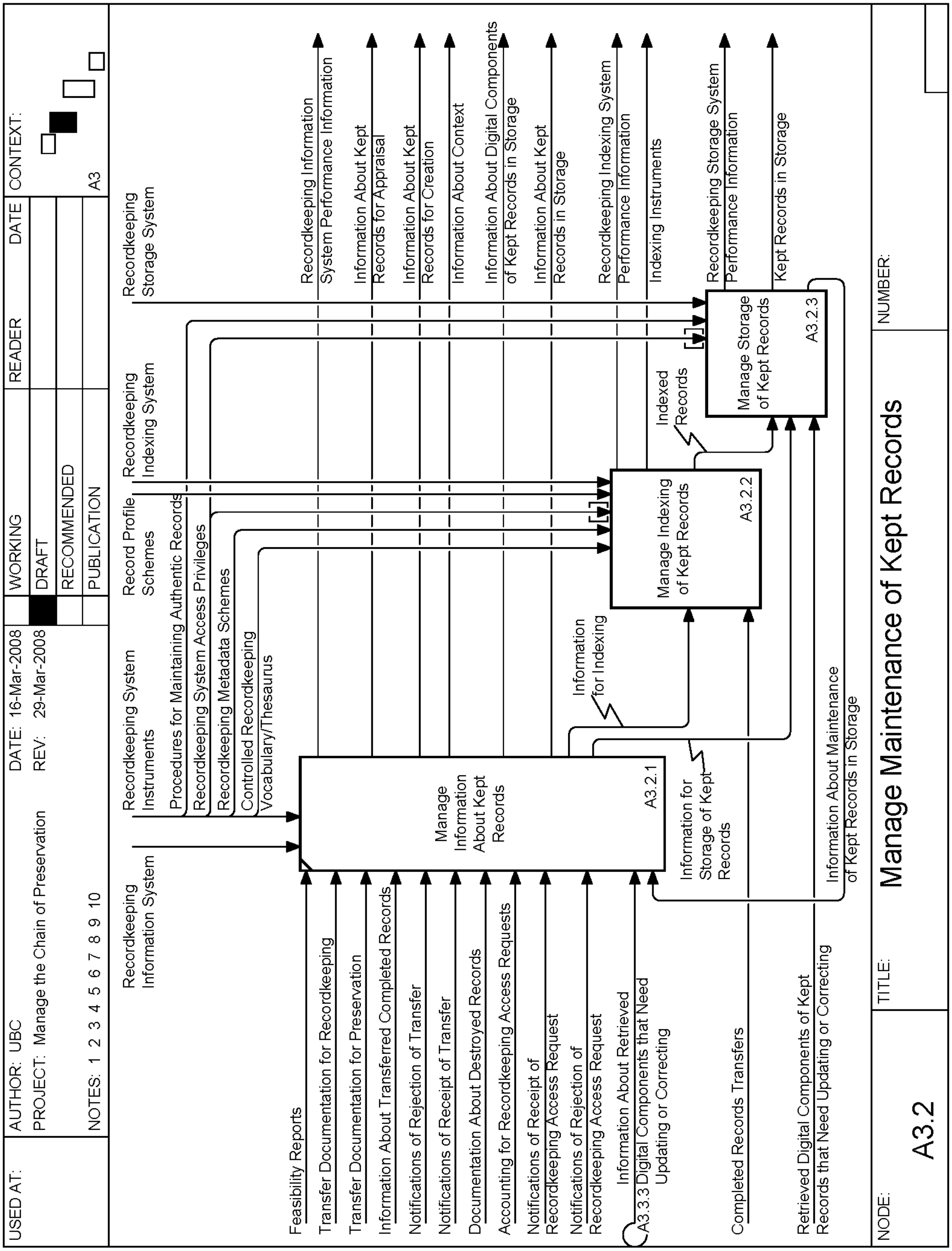

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
DATE: 10-Mar-2008
REV: 24-Mar-2008
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.2.3
Design Requirements
Information About Available Technology
Determine Preservation System Performance Requirements
A1.2.3.2.1
Preservation Information System Performance Requirements
Preservation Selection System Performance Requirements
Preservation Acquisition System Performance Requirements
Preservation Description System Performance Requirements
Preservation Storage System Performance Requirements
Preservation Retrieval System Performance Requirements
Preservation Access System Performance Requirements
Preservation System Performance Requirements
Determine Preservation System Monitoring Requirements
A1.2.3.2.2
Preservation Information System Monitoring Requirements
Preservation Selection System Monitoring Requirements
Preservation Acquisition System Monitoring Requirements
Preservation Description System Monitoring Requirements
Preservation Storage System Monitoring Requirements
Preservation Retrieval System Monitoring Requirements
Preservation Access System Monitoring Requirements
Preservation System Monitoring Requirements
Determine Preservation System Technological Requirements
A1.2.3.2.3
Preservation Information System Technological Requirements
Preservation Selection System Technological Requirements
Preservation Acquisition System Technological Requirements
Preservation Description System Technological Requirements
Preservation Storage System Technological Requirements
Preservation Retrieval System Technological Requirements
Preservation Access System Technological Requirements
Preservation System Technological Requirements
NODE:
A1.2.3.2
TITLE:
Establish Preservation System Functional Requirements
NUMBER:

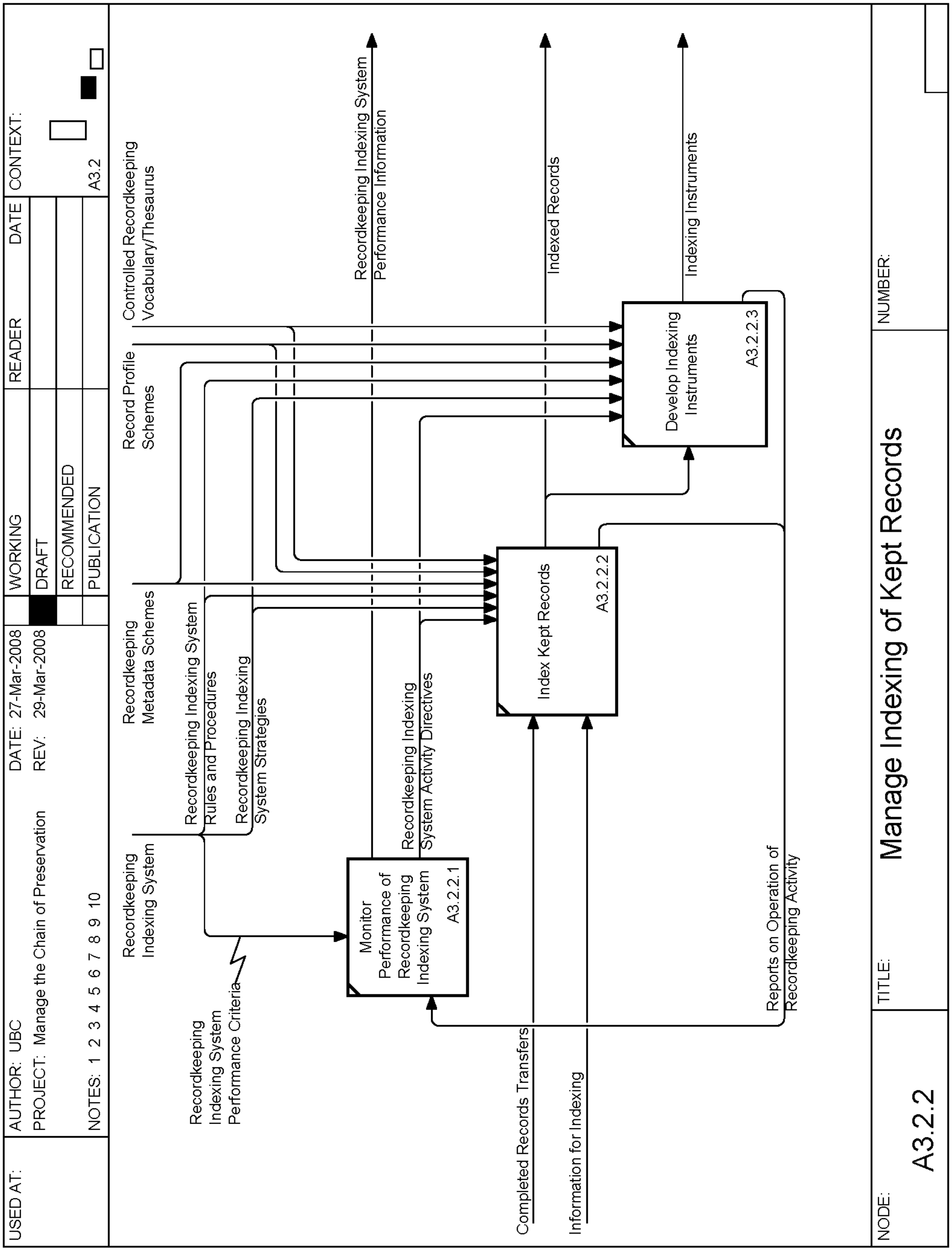

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 19-Jan-2004
REV: 06-Apr-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A0
Monitor Performance of Record-making System
A2.1
Manage Making and Receipt of Records
A2.2
Manage Setting Aside of Completed Records
A2.3
Creator's Existing Records
Information About Creator's Existing Records
Received Documents
Information About Received Documents' Context
Information About Kept Records for Creation
Record-making System Performance Criteria
Recordkeeping System
Integrated Business and Documentary Procedures
Procedures for Ensuring the Reliability of Records
Records Forms
Record-making System
Procedures for Maintaining Authentic Records
Record-making Metadata Schemes
Record Profile Schemes
Procedures for Ensuring the Accuracy of Records
Record-making Activity Directives
Completed Records
Information About Executed Records
Information About Outgoing Documents
Record Copies of Outgoing Documents
Record-making System Performance Information
Outgoing Documents
Information About Transferred Completed Records
Transfer Documentation for Recordkeeping
Completed Records Transfers
Record-making Capture System Performance Information
Record-making Identification System Performance Information
Record-making Declaration System Performance Information
Record-making Execution System Performance Information
Record-making Transfer System Performance Information
Creator
Records Manager
NODE:
A2
TITLE:
Manage Records in a Record-making System
NUMBER:

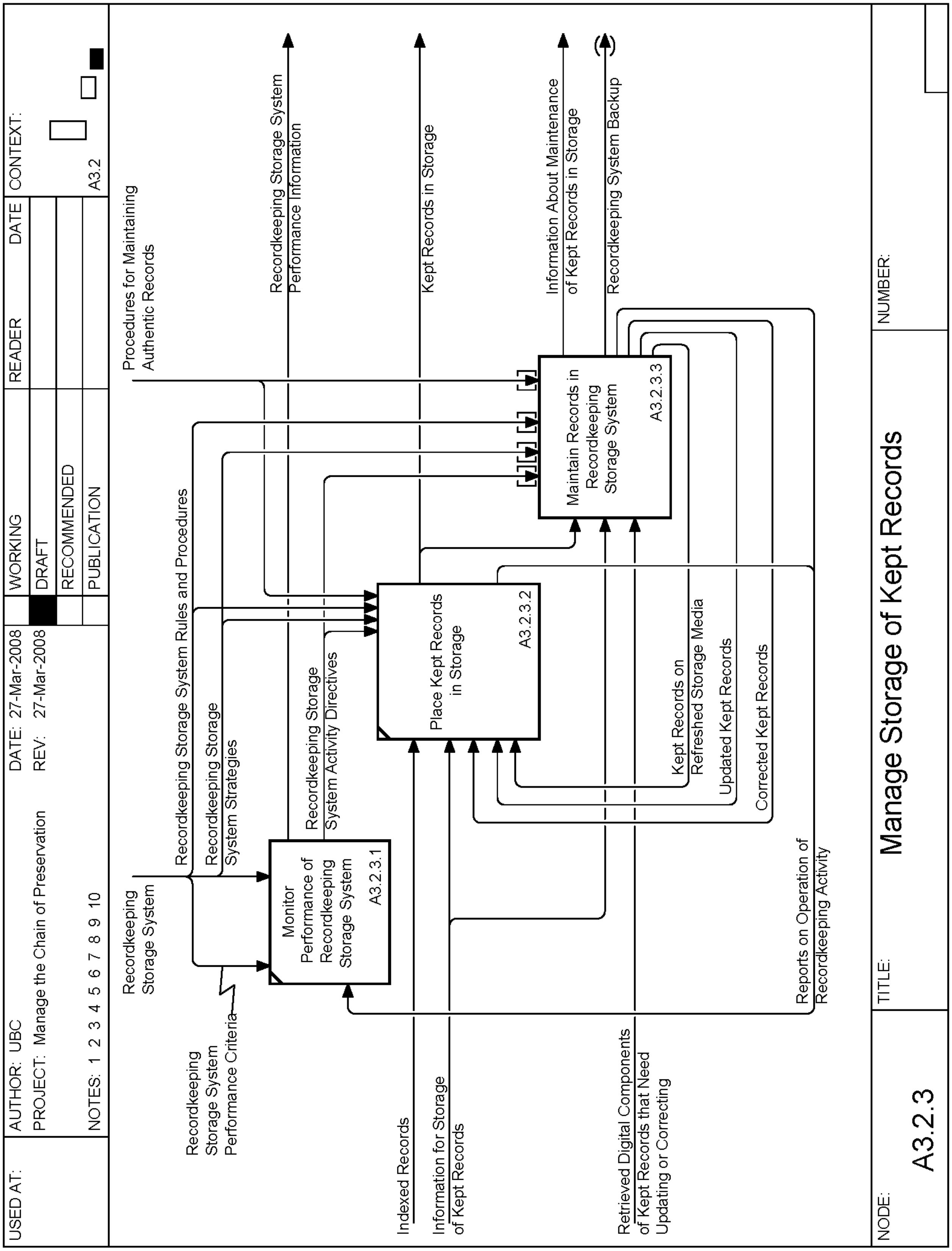

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Mar-2004
REV: 31-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A2
Records
Forms
Procedures for Ensuring the Reliability of Records
Information About Kept Records for Creation
Make Documents
A2.2.1
Made Documents
Capture Documents
A2.2.2
Received Documents
Information About Made Documents' Context
Captured Documents
Recordkeeping Classification Scheme
Recordkeeping System
Recordkeeping Registration Scheme
Identify Documents
A2.2.3
Information About Received Documents' Context
Information About Creator's Existing Records
Creator's Existing Records
Identified Documents
Declare Records
A2.2.4
Declared Records
Execute Records
A2.2.5
Record-making Capture System Performance Information
Record-making Identification System Performance Information
Record-making Declaration System Performance Information
Record-making Execution System Performance Information
Outgoing Documents
Information About Outgoing Documents
Record Copies of Outgoing Documents
Completed Records
Information About Executed Records
NODE:
A2.2
TITLE:
Manage Making and Receipt of Records
NUMBER:

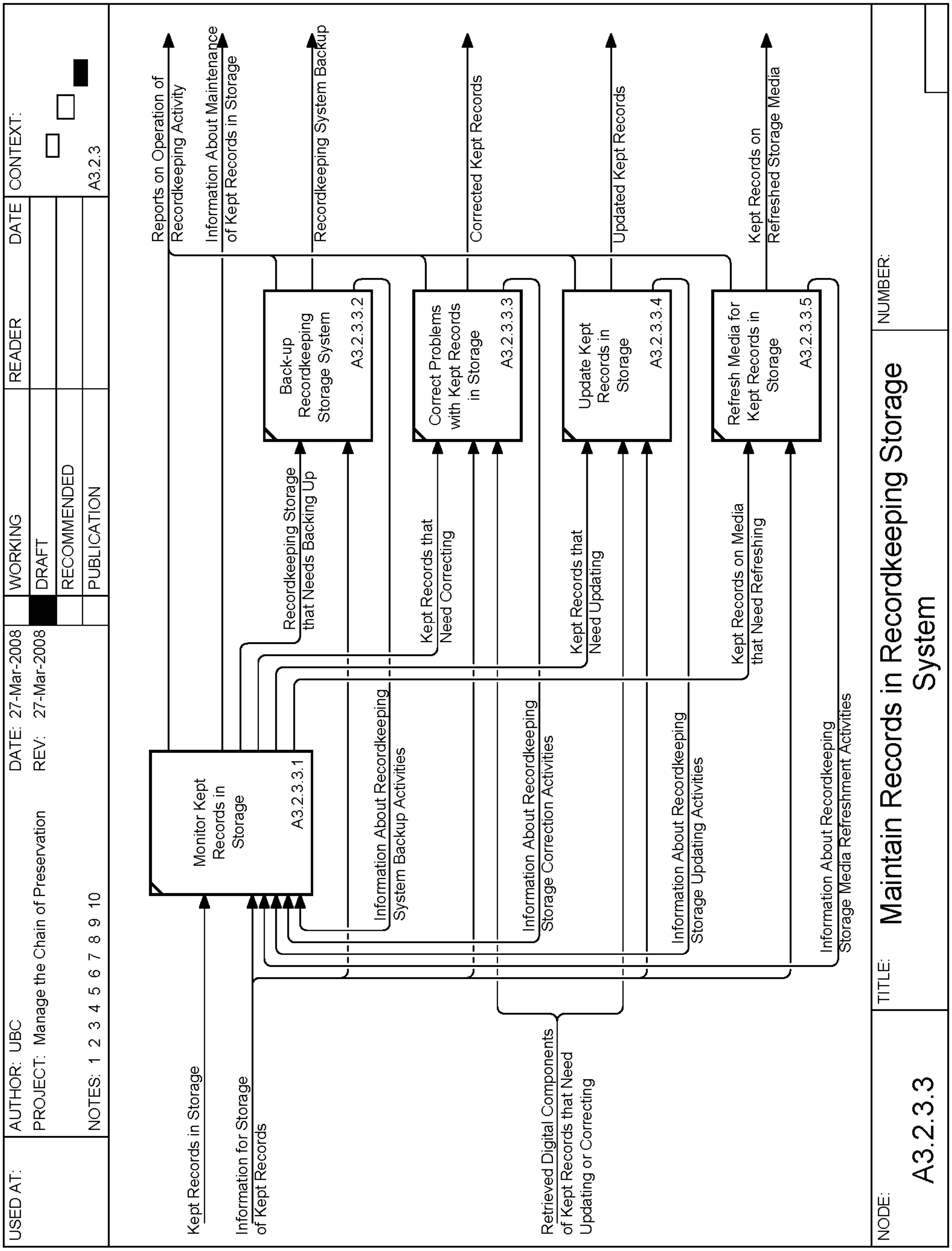

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
DATE: 20-Mar-2008
REV: 28-Mar-2008
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A2
Record-making Transfer System Performance Criteria
Record-making Metadata Schemes
Record Profile Schemes
Record-making Transfer System Performance Information
Monitor Performance of Record-making Transfer System
A2.3.1
Record-making Transfer System Activity Directives
Prepare Completed Records for Transfer to Recordkeeping System
A2.3.2
Completed Records Prepared for Transfer to Recordkeeping System
Information About Completed Records Prepared for Transfer to Recordkeeping System
Transfer Completed Records to Recordkeeping System
A2.3.3
Completed Records Transfers
Information About Transferred Completed Records
Transfer Documentation for Recordkeeping
Completed Records
Record Copies of Outgoing Documents
Information About Executed Records
Information About Outgoing Documents
Reports on Operation of Record-making Activity
NODE:
A2.3
TITLE:
Manage Setting Aside of Completed Records
NUMBER:

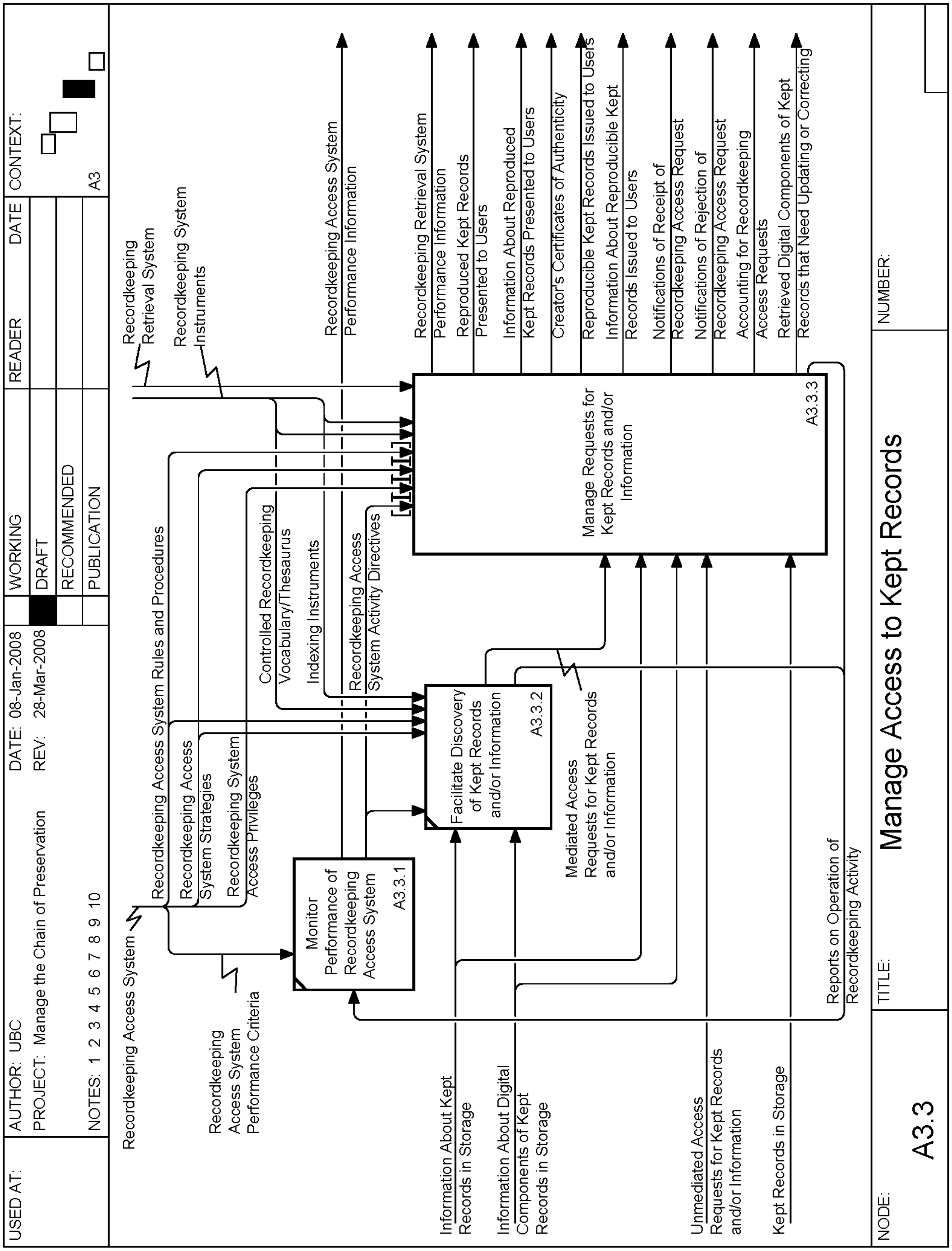

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 19-Jan-2004
REV: 28-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A0
Recordkeeping Indexing System
Recordkeeping System
Record Profile Schemes
Recordkeeping Information System
Recordkeeping Storage System
Recordkeeping Access System
Recordkeeping Activity Directives
Recordkeeping System Instruments
Recordkeeping Retrieval System
Disposition Rules and Procedures
Terms and Conditions of Transfer
Recordkeeping Disposition System
Recordkeeping System Performance Criteria
Recordkeeping Disposition System Performance Information
Recordkeeping Access System Performance Information
Feasibility Reports
Completed Records Transfers
Transfer Documentation for Recordkeeping
Information About Transferred Completed Records
Retrieved Digital Components of Kept Records that Need Updating or Correcting
Recordkeeping Retrieval System Performance Information
Unmediated Access Requests for Kept Records and/or Information
Notifications of Receipt of Transfer
Notifications of Rejection of Transfer
Monitor Performance of Recordkeeping System
A3.1
Manage Maintenance of Kept Records
A3.2
Manage Access to Kept Records
A3.3
Manage Disposition of Kept Records
A3.4
Information About Kept Records in Storage
Kept Records in Storage
Recordkeeping Information System Performance Information
Recordkeeping Indexing System Performance Information
Recordkeeping Storage System Performance Information
Information About Digital Components of Kept Records in Storage
Accounting for Recordkeeping Access Requests
Recordkeeping System Performance Information
Information About Kept Records for Creation
Information About Context
Information About Kept Records for Appraisal
Indexing Instruments
Reproduced Kept Records Presented to Users
Information About Reproduced Kept Records Presented to Users
Creator's Certificates of Authenticity
Reproducible Kept Records Issued to Users
Information About Reproducible Kept Records Issued to Users
Notifications of Rejection of Recordkeeping Access Request
Notifications of Receipt of Recordkeeping Access Request
Kept Records Transfers
Information About Transferred Kept Records
Transfer Documentation for Preservation
Documentation About Destroyed Records
NODE:
A3
TITLE:
Manage Records in a Recordkeeping System
NUMBER:

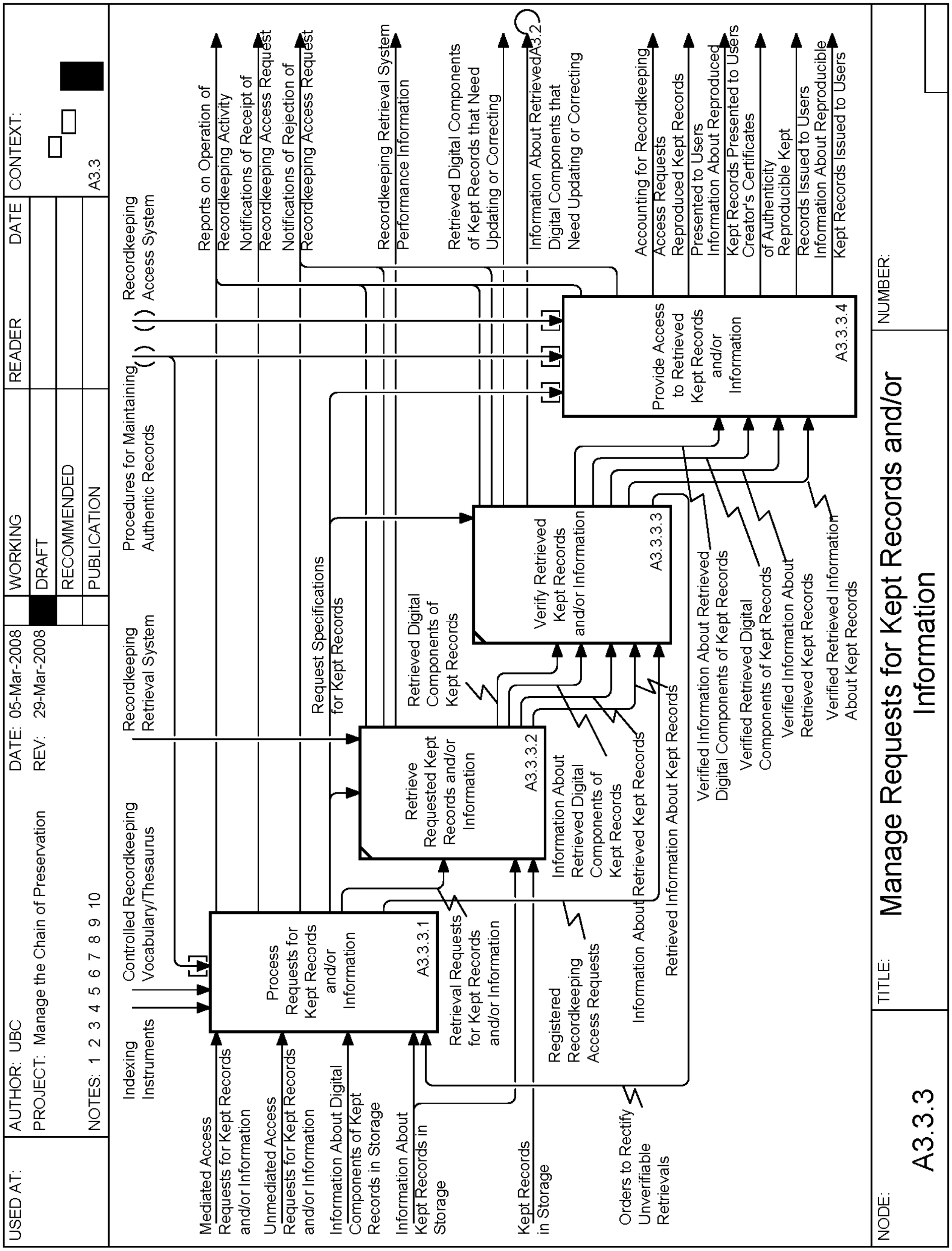

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
DATE: 16-Mar-2008
REV: 29-Mar-2008
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3
Recordkeeping Information System
Recordkeeping System Instruments
Procedures for Maintaining Authentic Records
Recordkeeping System Access Privileges
Recordkeeping Metadata Schemes
Controlled Recordkeeping Vocabulary/Thesaurus
Record Profile Schemes
Recordkeeping Indexing System
Recordkeeping Storage System
Feasibility Reports
Transfer Documentation for Recordkeeping
Transfer Documentation for Preservation
Information About Transferred Completed Records
Notifications of Rejection of Transfer
Notifications of Receipt of Transfer
Documentation About Destroyed Records
Accounting for Recordkeeping Access Requests
Notifications of Receipt of Recordkeeping Access Request
Notifications of Rejection of Recordkeeping Access Request
Information About Retrieved
A3.3.3 Digital Components that Need Updating or Correcting
Completed Records Transfers
Retrieved Digital Components of Kept Records that Need Updating or Correcting
Manage Information About Kept Records
A3.2.1
Information for Indexing
Information for Storage of Kept Records
Information About Maintenance of Kept Records in Storage
Manage Indexing of Kept Records
A3.2.2
Indexed Records
Manage Storage of Kept Records
A3.2.3
Recordkeeping Information System Performance Information
Information About Kept Records for Appraisal
Information About Kept Records for Creation
Information About Context
Information About Digital Components of Kept Records in Storage
Information About Kept Records in Storage
Recordkeeping Indexing System Performance Information
Indexing Instruments
Recordkeeping Storage System Performance Information
Kept Records in Storage
NODE:
A3.2
TITLE:
Manage Maintenance of Kept Records
NUMBER:

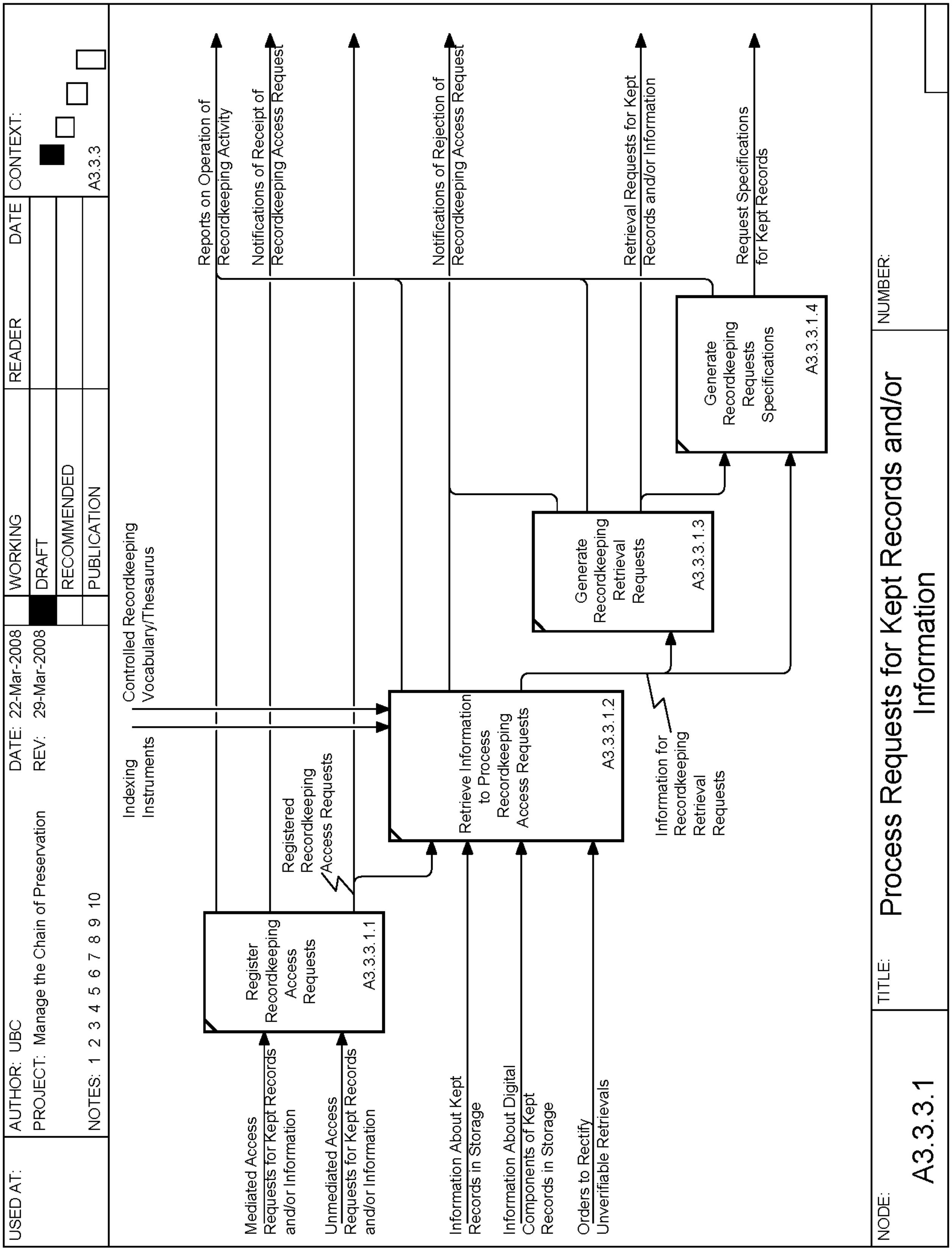

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
DATE: 27-Mar-2008
REV: 29-Mar-2008
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3.2
Recordkeeping Indexing System
Recordkeeping Indexing System Performance Criteria
Recordkeeping Metadata Schemes
Recordkeeping Indexing System Rules and Procedures
Recordkeeping Indexing System Strategies
Record Profile Schemes
Controlled Recordkeeping Vocabulary/Thesaurus
Recordkeeping Indexing System Performance Information
Recordkeeping Indexing System Activity Directives
Indexed Records
Indexing Instruments
Monitor Performance of Recordkeeping Indexing System
A3.2.2.1
Index Kept Records
A3.2.2.2
Develop Indexing Instruments
A3.2.2.3
Completed Records Transfers
Information for Indexing
Reports on Operation of Recordkeeping Activity
NODE:
A3.2.2
TITLE:
Manage Indexing of Kept Records
NUMBER:

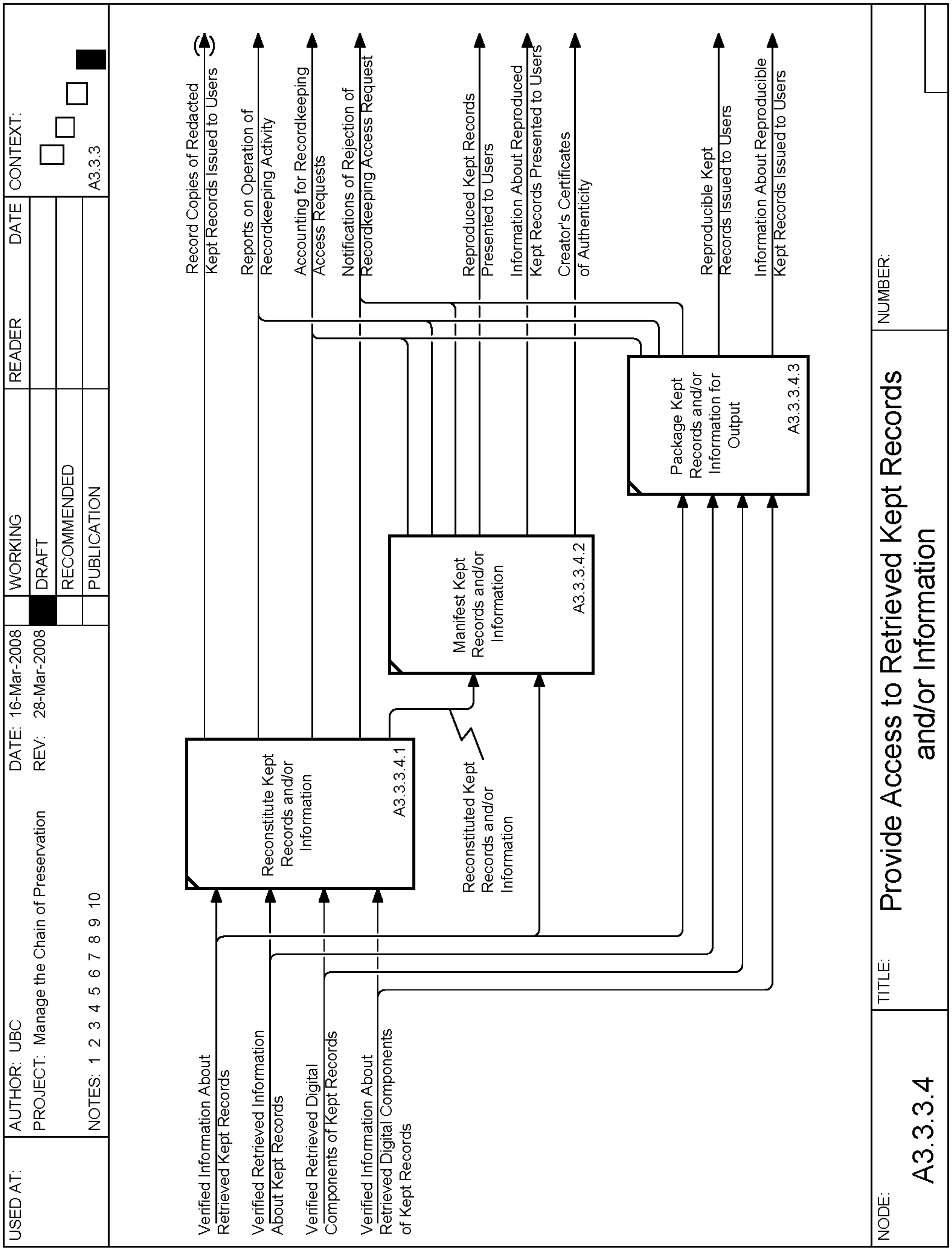

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
DATE: 27-Mar-2008
REV: 27-Mar-2008
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3.2
Recordkeeping Storage System
Recordkeeping Storage System Performance Criteria
Indexed Records
Information for Storage of Kept Records
Retrieved Digital Components of Kept Records that Need Updating or Correcting
Recordkeeping Storage System Rules and Procedures
Recordkeeping Storage System Strategies
Recordkeeping Storage System Activity Directives
Procedures for Maintaining Authentic Records
Monitor Performance of Recordkeeping Storage System
A3.2.3.1
Place Kept Records in Storage
A3.2.3.2
Maintain Records in Recordkeeping Storage System
A3.2.3.3
Kept Records on Refreshed Storage Media
Updated Kept Records
Corrected Kept Records
Reports on Operation of Recordkeeping Activity
Recordkeeping Storage System Performance Information
Kept Records in Storage
Information About Maintenance of Kept Records in Storage
Recordkeeping System Backup
NODE:
A3.2.3
TITLE:
Manage Storage of Kept Records
NUMBER:

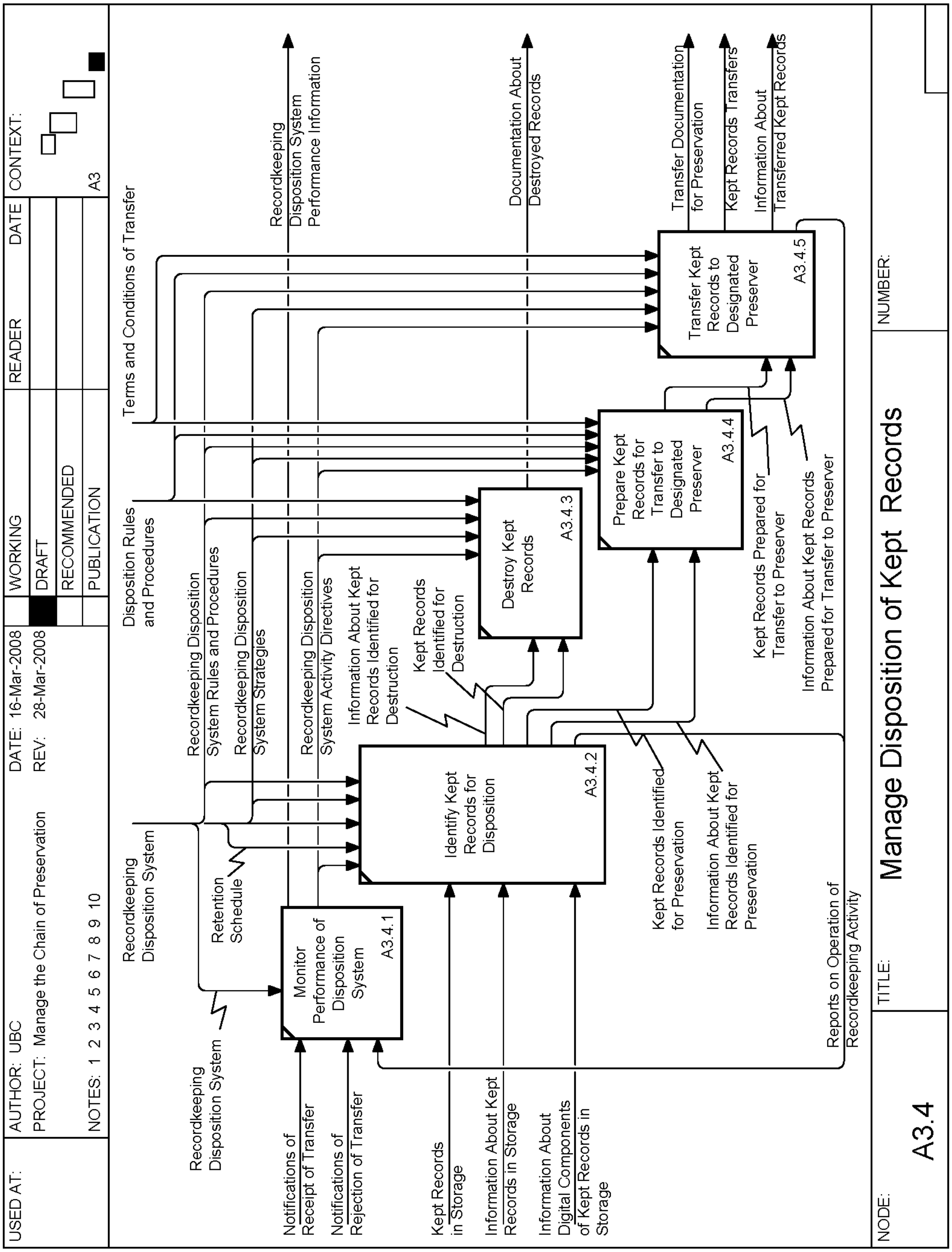

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
DATE: 27-Mar-2008
REV: 27-Mar-2008
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3.2.3
Kept Records in Storage
Information for Storage of Kept Records
Retrieved Digital Components of Kept Records that Need Updating or Correcting
Monitor Kept Records in Storage
A3.2.3.3.1
Reports on Operation of Recordkeeping Activity
Information About Maintenance of Kept Records in Storage
Recordkeeping Storage that Needs Backing Up
Back-up Recordkeeping Storage System
A3.2.3.3.2
Recordkeeping System Backup
Information About Recordkeeping System Backup Activities
Kept Records that Need Correcting
Correct Problems with Kept Records in Storage
A3.2.3.3.3
Corrected Kept Records
Information About Recordkeeping Storage Correction Activities
Kept Records that Need Updating
Update Kept Records in Storage
A3.2.3.3.4
Updated Kept Records
Information About Recordkeeping Storage Updating Activities
Kept Records on Media that Need Refreshing
Refresh Media for Kept Records in Storage
A3.2.3.3.5
Kept Records on Refreshed Storage Media
Information About Recordkeeping Storage Media Refreshment Activities
NODE:
A3.2.3.3
TITLE:
Maintain Records in Recordkeeping Storage System
NUMBER:

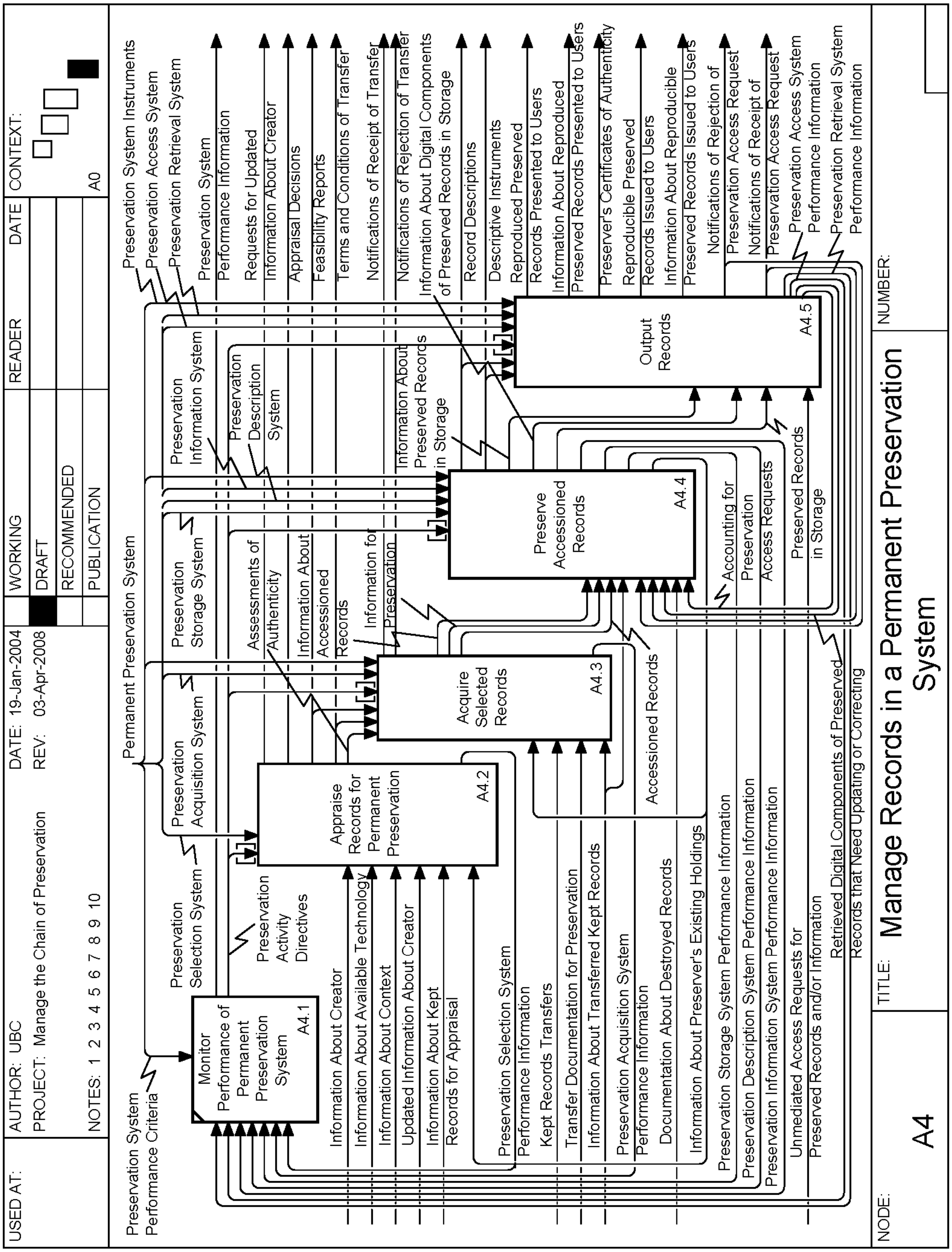

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Jan-2008
REV: 28-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3
Recordkeeping Access System
Recordkeeping Access System Performance Criteria
Recordkeeping Access System Rules and Procedures
Recordkeeping Access System Strategies
Recordkeeping System Access Privileges
Controlled Recordkeeping Vocabulary/Thesaurus
Indexing Instruments
Recordkeeping Access System Activity Directives
Recordkeeping Retrieval System
Recordkeeping System Instruments
Monitor Performance of Recordkeeping Access System
A3.3.1
Facilitate Discovery of Kept Records and/or Information
A3.3.2
Manage Requests for Kept Records and/or Information
A3.3.3
Information About Kept Records in Storage
Information About Digital Components of Kept Records in Storage
Unmediated Access Requests for Kept Records and/or Information
Kept Records in Storage
Mediated Access Requests for Kept Records and/or Information
Reports on Operation of Recordkeeping Activity
Recordkeeping Access System Performance Information
Recordkeeping Retrieval System Performance Information
Reproduced Kept Records Presented to Users
Information About Reproduced Kept Records Presented to Users
Creator's Certificates of Authenticity
Reproducible Kept Records Issued to Users
Information About Reproducible Kept Records Issued to Users
Notifications of Receipt of Recordkeeping Access Request
Notifications of Rejection of Recordkeeping Access Request
Accounting for Recordkeeping Access Requests
Retrieved Digital Components of Kept Records that Need Updating or Correcting
NODE:
A3.3
TITLE:
Manage Access to Kept Records
NUMBER:

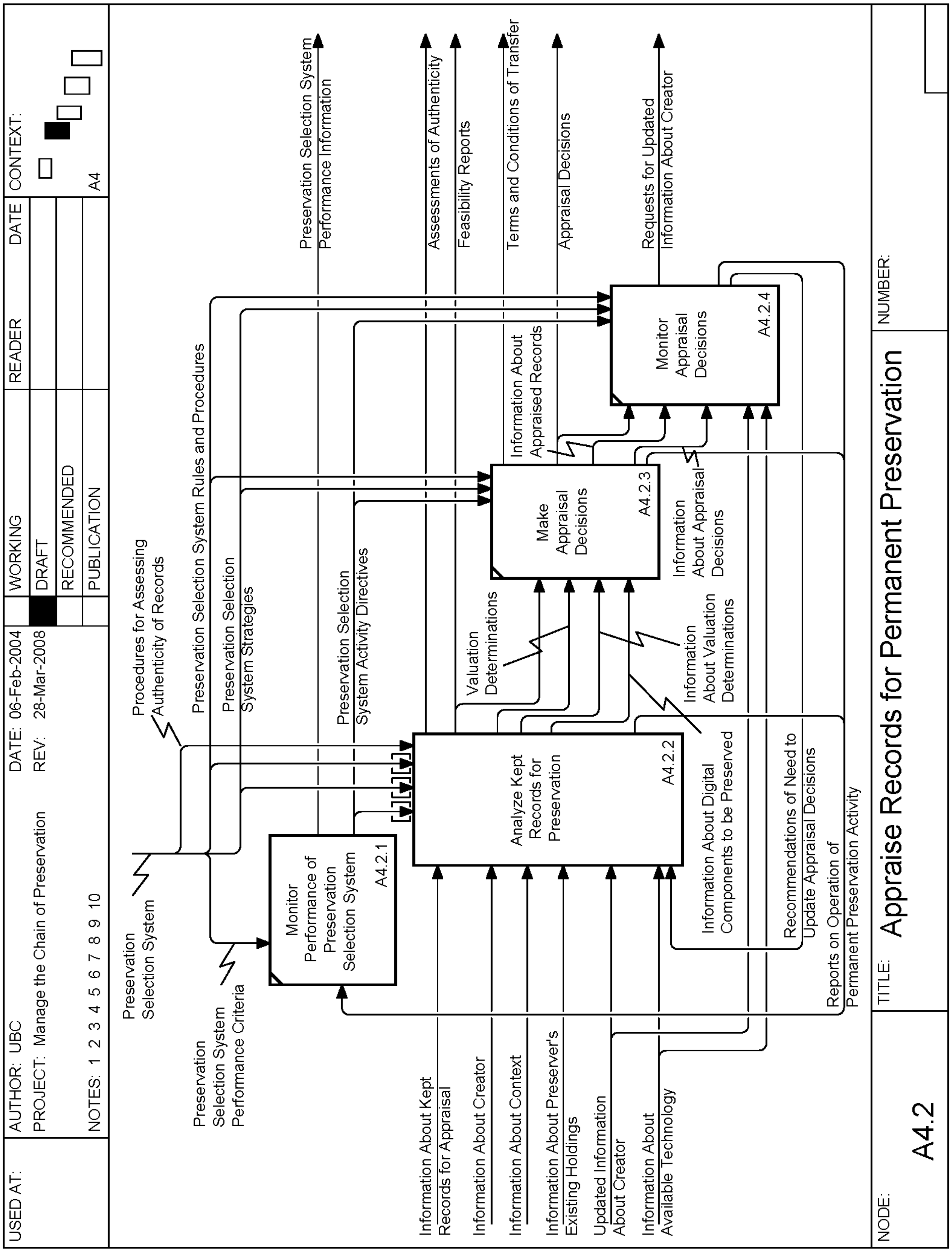

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 05-Mar-2008
REV: 29-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3.3
Indexing Instruments
Controlled Recordkeeping Vocabulary/Thesaurus
Recordkeeping Retrieval System
Procedures for Maintaining Authentic Records
Recordkeeping Access System
Mediated Access Requests for Kept Records and/or Information
Unmediated Access Requests for Kept Records and/or Information
Information About Digital Components of Kept Records in Storage
Information About Kept Records in Storage
Kept Records in Storage
Orders to Rectify Unverifiable Retrievals
Process Requests for Kept Records and/or Information
A3.3.3.1
Retrieval Requests for Kept Records and/or Information
Registered Recordkeeping Access Requests
Information About Retrieved Kept Records
Retrieved Information About Kept Records
Request Specifications for Kept Records
Retrieve Requested Kept Records and/or Information
A3.3.3.2
Information About Retrieved Digital Components of Kept Records
Retrieved Digital Components of Kept Records
Verify Retrieved Kept Records and/or Information
A3.3.3.3
Verified Information About Retrieved Digital Components of Kept Records
Verified Retrieved Digital Components of Kept Records
Verified Information About Retrieved Kept Records
Verified Retrieved Information About Kept Records
Provide Access to Retrieved Kept Records and/or Information
A3.3.3.4
Reports on Operation of Recordkeeping Activity
Notifications of Receipt of Recordkeeping Access Request
Notifications of Rejection of Recordkeeping Access Request
Recordkeeping Retrieval System Performance Information
Retrieved Digital Components of Kept Records that Need Updating or Correcting
Information About Retrieved Digital Components that Need Updating or Correcting
A3.2
Accounting for Recordkeeping Access Requests
Reproduced Kept Records Presented to Users
Information About Reproduced Kept Records Presented to Users
Creator's Certificates of Authenticity
Reproducible Kept Records Issued to Users
Information About Reproducible Kept Records Issued to Users
NODE:
A3.3.3
TITLE:
Manage Requests for Kept Records and/or Information
NUMBER:

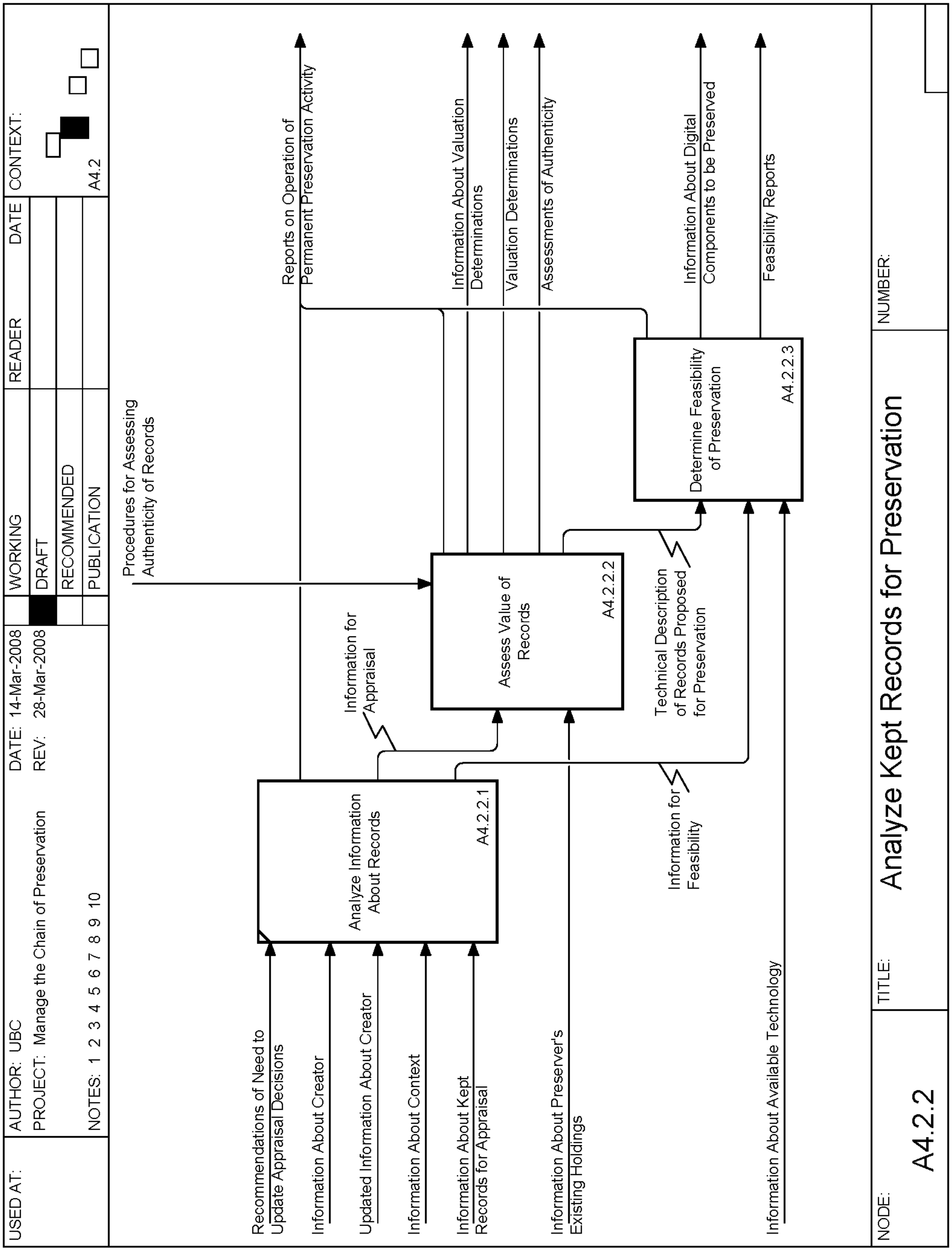

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 22-Mar-2008
REV: 29-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3.3.3
Indexing Instruments
Controlled Recordkeeping Vocabulary/Thesaurus
Mediated Access Requests for Kept Records and/or Information
Unmediated Access Requests for Kept Records and/or Information
Information About Kept Records in Storage
Information About Digital Components of Kept Records in Storage
Orders to Rectify Unverifiable Retrievals
Register Recordkeeping Access Requests
A3.3.3.1.1
Registered Recordkeeping Access Requests
Retrieve Information to Process Recordkeeping Access Requests
A3.3.3.1.2
Information for Recordkeeping Retrieval Requests
Generate Recordkeeping Retrieval Requests
A3.3.3.1.3
Generate Recordkeeping Requests Specifications
A3.3.3.1.4
Reports on Operation of Recordkeeping Activity
Notifications of Receipt of Recordkeeping Access Request
Notifications of Rejection of Recordkeeping Access Request
Retrieval Requests for Kept Records and/or Information
Request Specifications for Kept Records
NODE:
A3.3.3.1
TITLE:
Process Requests for Kept Records and/or Information
NUMBER:

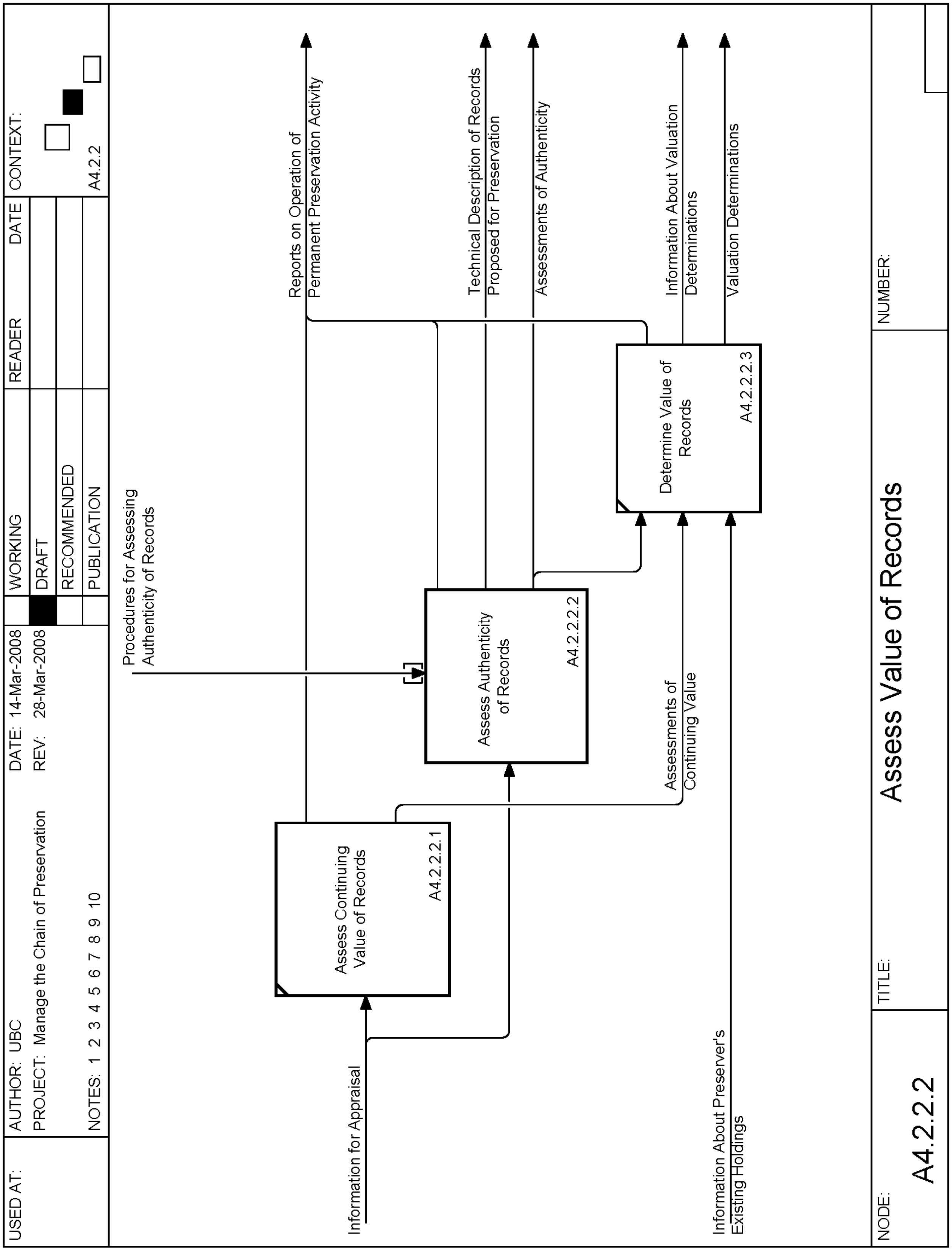

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 16-Mar-2008
REV: 28-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3.3.3
Verified Information About Retrieved Kept Records
Verified Retrieved Information About Kept Records
Verified Retrieved Digital Components of Kept Records
Verified Information About Retrieved Digital Components of Kept Records
Reconstitute Kept Records and/or Information
A3.3.3.4.1
Reconstituted Kept Records and/or Information
Manifest Kept Records and/or Information
A3.3.3.4.2
Package Kept Records and/or Information for Output
A3.3.3.4.3
Record Copies of Redacted Kept Records Issued to Users
Reports on Operation of Recordkeeping Activity
Accounting for Recordkeeping Access Requests
Notifications of Rejection of Recordkeeping Access Request
Reproduced Kept Records Presented to Users
Information About Reproduced Kept Records Presented to Users
Creator's Certificates of Authenticity
Reproducible Kept Records Issued to Users
Information About Reproducible Kept Records Issued to Users
NODE:
A3.3.3.4
TITLE:
Provide Access to Retrieved Kept Records and/or Information
NUMBER:

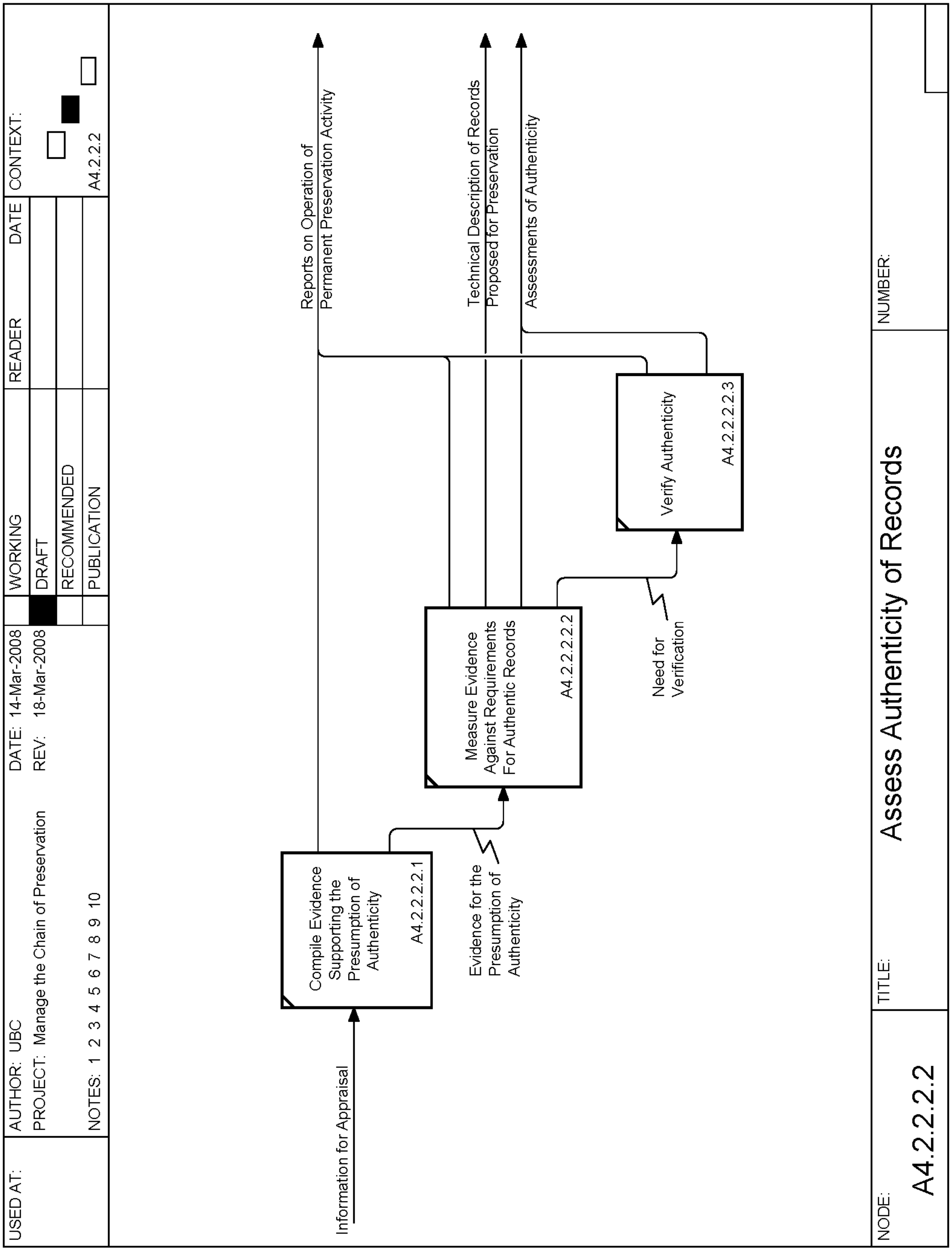

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 16-Mar-2008
REV: 28-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3
Notifications of Receipt of Transfer
Notifications of Rejection of Transfer
Kept Records in Storage
Information About Kept Records in Storage
Information About Digital Components of Kept Records in Storage
Recordkeeping Disposition System
Recordkeeping Disposition System
Retention Schedule
Monitor Performance of Disposition System
A3.4.1
Identify Kept Records for Disposition
A3.4.2
Destroy Kept Records
A3.4.3
Prepare Kept Records for Transfer to Designated Preserver
A3.4.4
Transfer Kept Records to Designated Preserver
A3.4.5
Disposition Rules and Procedures
Terms and Conditions of Transfer
Recordkeeping Disposition System Rules and Procedures
Recordkeeping Disposition System Strategies
Recordkeeping Disposition System Activity Directives
Information About Kept Records Identified for Destruction
Kept Records Identified for Destruction
Recordkeeping Disposition System Performance Information
Documentation About Destroyed Records
Transfer Documentation for Preservation
Kept Records Transfers
Information About Transferred Kept Records
Kept Records Identified for Preservation
Information About Kept Records Identified for Preservation
Kept Records Prepared for Transfer to Preserver
Information About Kept Records Prepared for Transfer to Preserver
Reports on Operation of Recordkeeping Activity
NODE:
A3.4
TITLE:
Manage Disposition of Kept Records
NUMBER:

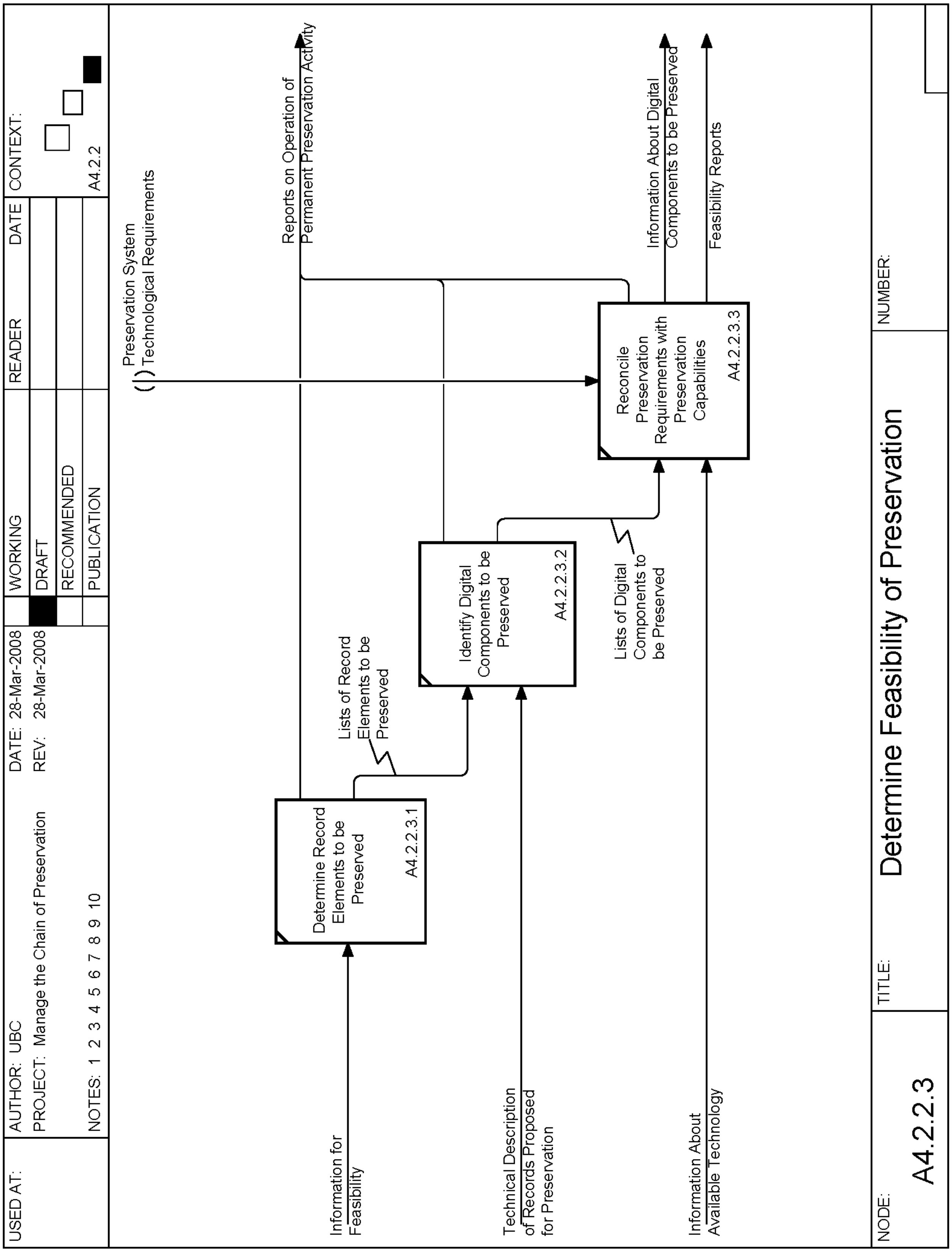

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 19-Jan-2004
REV: 03-Apr-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A0
Preservation System Performance Criteria
Permanent Preservation System
Preservation Selection System
Preservation Activity Directives
Preservation Acquisition System
Preservation Storage System
Preservation Information System
Preservation Description System
Preservation System Instruments
Preservation Access System
Preservation Retrieval System
Monitor Performance of Permanent Preservation System
A4.1
Appraise Records for Permanent Preservation
A4.2
Acquire Selected Records
A4.3
Preserve Accessioned Records
A4.4
Output Records
A4.5
Information About Creator
Information About Available Technology
Information About Context
Updated Information About Creator
Information About Kept Records for Appraisal
Preservation Selection System Performance Information
Kept Records Transfers
Transfer Documentation for Preservation
Information About Transferred Kept Records
Preservation Acquisition System Performance Information
Documentation About Destroyed Records
Information About Preserver's Existing Holdings
Preservation Storage System Performance Information
Preservation Description System Performance Information
Preservation Information System Performance Information
Unmediated Access Requests for Preserved Records and/or Information
Retrieved Digital Components of Preserved Records that Need Updating or Correcting
Assessments of Authenticity
Information About Accessioned Records
Information for Preservation
Accessioned Records
Information About Preserved Records in Storage
Accounting for Preservation Access Requests
Preserved Records in Storage
Preservation System Performance Information
Requests for Updated Information About Creator
Appraisal Decisions
Feasibility Reports
Terms and Conditions of Transfer
Notifications of Receipt of Transfer
Notifications of Rejection of Transfer
Information About Digital Components of Preserved Records in Storage
Record Descriptions
Descriptive Instruments
Reproduced Preserved Records Presented to Users
Information About Reproduced Preserved Records Presented to Users
Preserver's Certificates of Authenticity
Reproducible Preserved Records Issued to Users
Information About Reproducible Preserved Records Issued to Users
Notifications of Rejection of Preservation Access Request
Notifications of Receipt of Preservation Access Request
Preservation Access System Performance Information
Preservation Retrieval System Performance Information
NODE:
A4
TITLE:
Manage Records in a Permanent Preservation System
NUMBER:

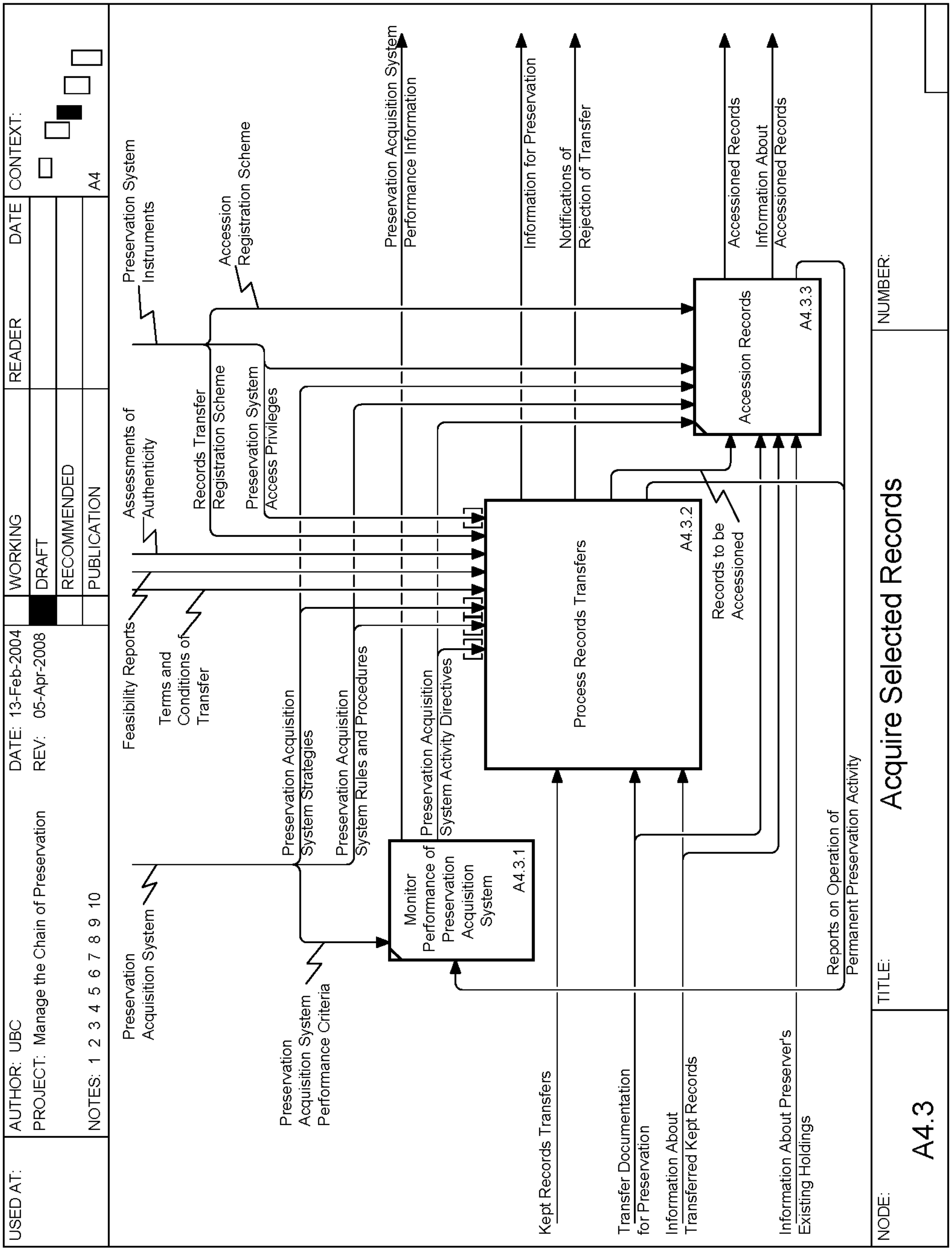

USED AT:
AUTHOR: UBC
DATE: 06-Feb-2004
REV: 28-Mar-2008
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4
Preservation Selection System
Preservation Selection System Performance Criteria
Procedures for Assessing Authenticity of Records
Preservation Selection System Rules and Procedures
Preservation Selection System Strategies
Preservation Selection System Activity Directives
Monitor Performance of Preservation Selection System
A4.2.1
Analyze Kept Records for Preservation
A4.2.2
Make Appraisal Decisions
A4.2.3
Monitor Appraisal Decisions
A4.2.4
Valuation Determinations
Information About Valuation Determinations
Information About Appraisal Decisions
Information About Appraised Records
Information About Kept Records for Appraisal
Information About Creator
Information About Context
Information About Preserver's Existing Holdings
Updated Information About Creator
Information About Available Technology
Information About Digital Components to be Preserved
Recommendations of Need to Update Appraisal Decisions
Reports on Operation of Permanent Preservation Activity
Preservation Selection System Performance Information
Assessments of Authenticity
Feasibility Reports
Terms and Conditions of Transfer
Appraisal Decisions
Requests for Updated Information About Creator
NODE:
A4.2
TITLE:
Appraise Records for Permanent Preservation
NUMBER:

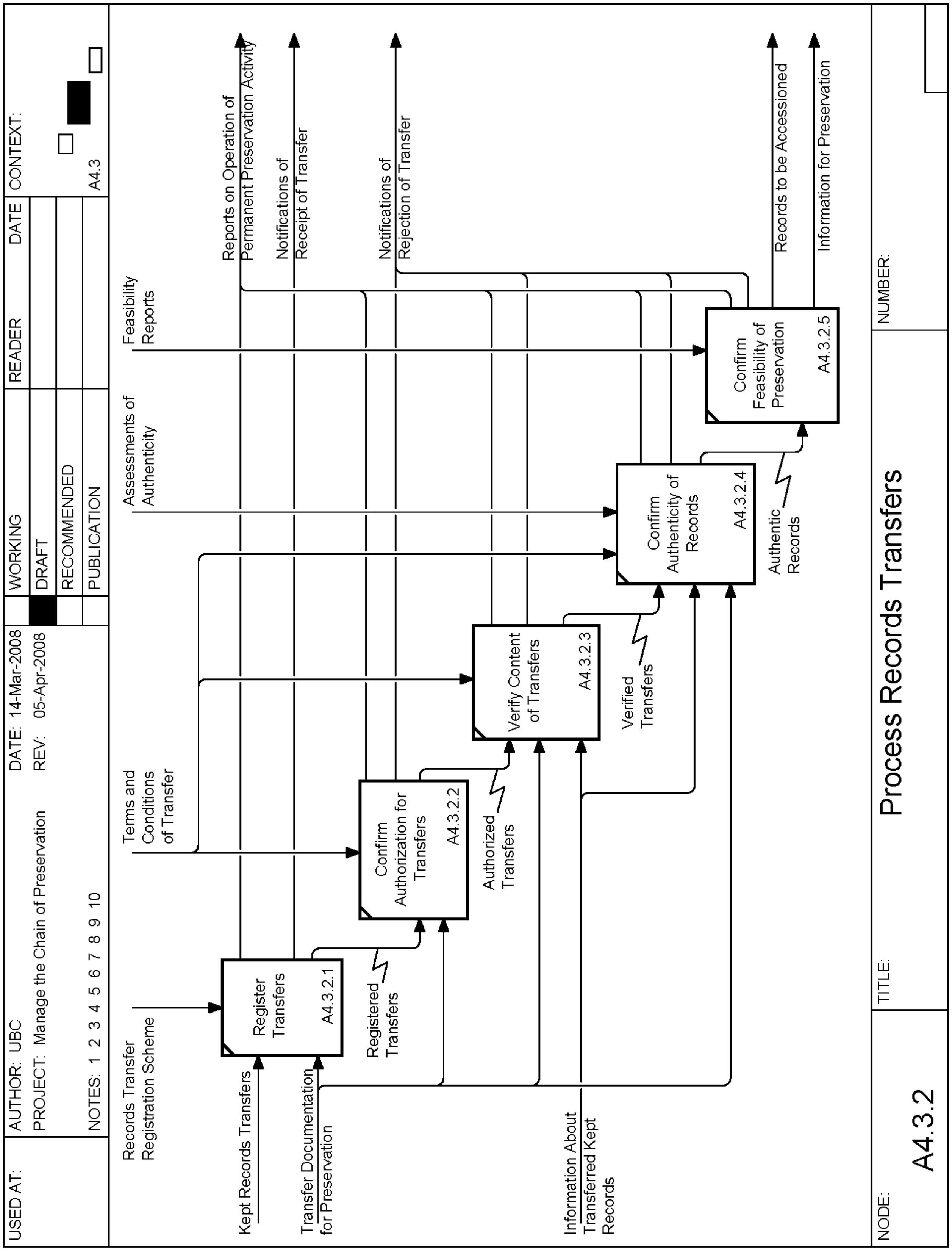

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 14-Mar-2008
REV: 28-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4.2
Recommendations of Need to Update Appraisal Decisions
Information About Creator
Updated Information About Creator
Information About Context
Information About Kept Records for Appraisal
Information About Preserver's Existing Holdings
Information About Available Technology
Procedures for Assessing Authenticity of Records
Analyze Information About Records
A4.2.2.1
Information for Appraisal
Information for Feasibility
Assess Value of Records
A4.2.2.2
Technical Description of Records Proposed for Preservation
Determine Feasibility of Preservation
A4.2.2.3
Reports on Operation of Permanent Preservation Activity
Information About Valuation Determinations
Valuation Determinations
Assessments of Authenticity
Information About Digital Components to be Preserved
Feasibility Reports
NODE:
A4.2.2
TITLE:
Analyze Kept Records for Preservation
NUMBER:

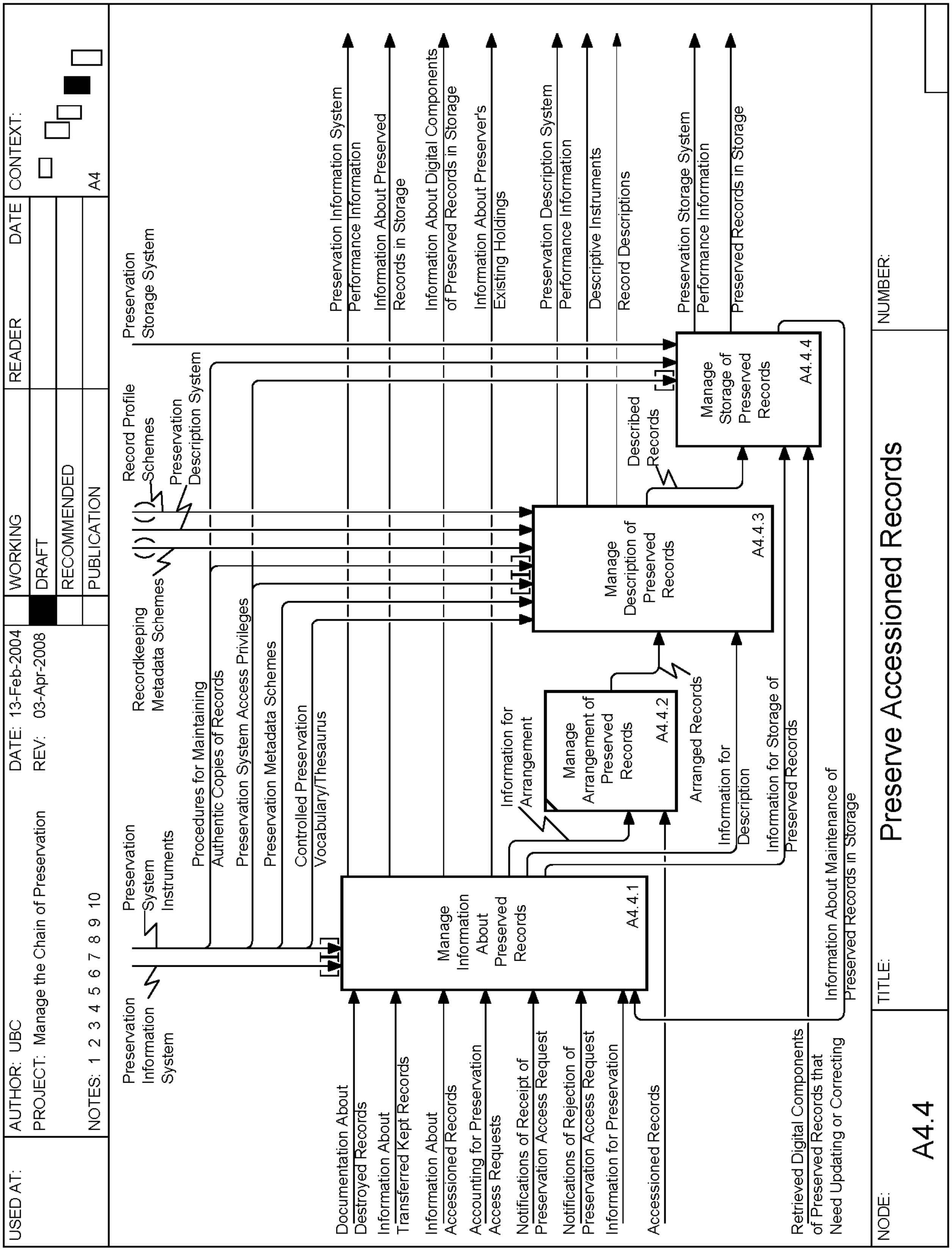

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 14-Mar-2008
REV: 28-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4.2.2
Information for Appraisal
Procedures for Assessing Authenticity of Records
Reports on Operation of Permanent Preservation Activity
Technical Description of Records Proposed for Preservation
Assessments of Authenticity
Information About Valuation Determinations
Valuation Determinations
Assess Continuing Value of Records
A4.2.2.2.1
Assess Authenticity of Records
A4.2.2.2.2
Determine Value of Records
A4.2.2.2.3
Assessments of Continuing Value
Information About Preserver's Existing Holdings
NODE:
A4.2.2.2
TITLE:
Assess Value of Records
NUMBER:

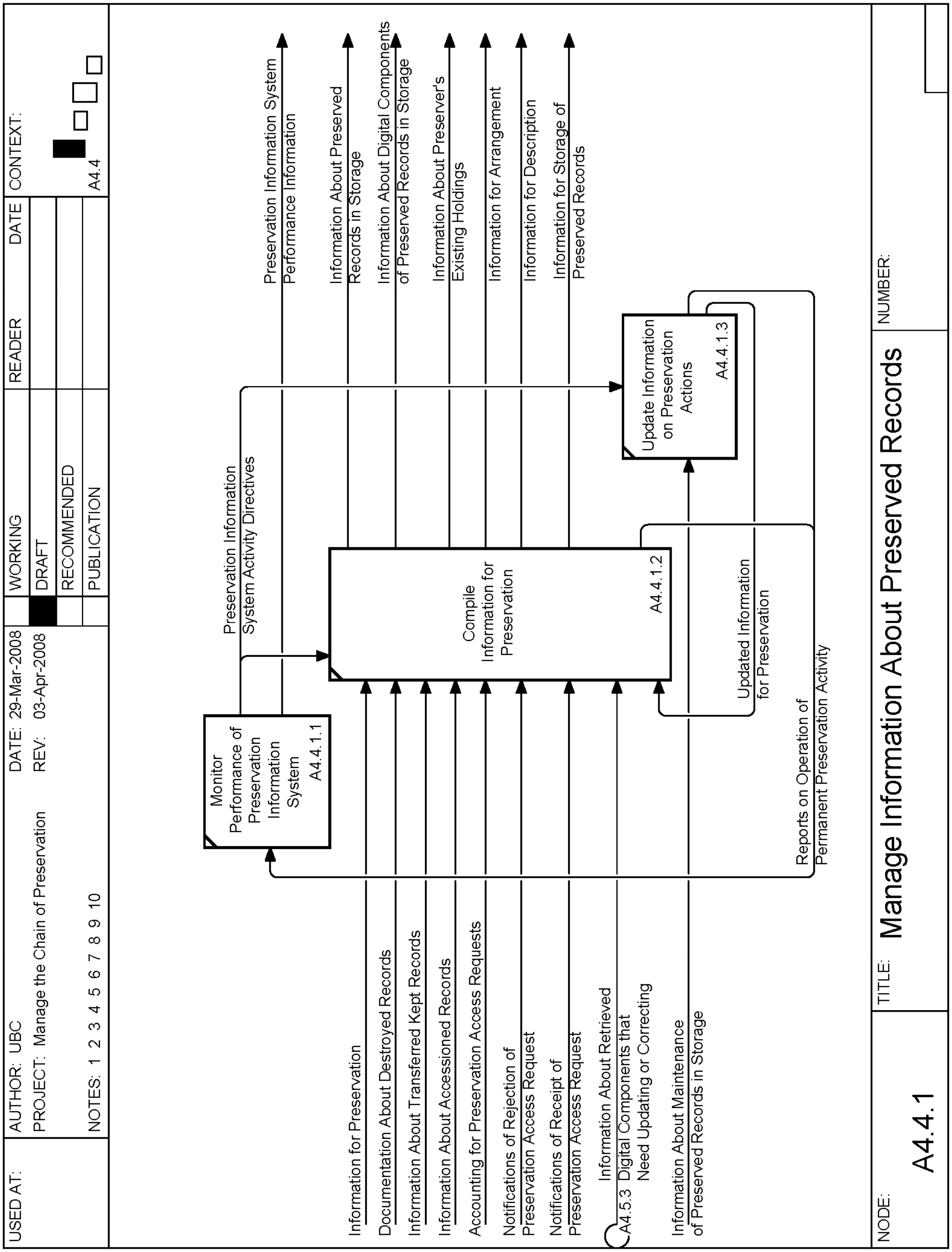

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 14-Mar-2008
REV: 18-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4.2.2.2
Information for Appraisal
Compile Evidence Supporting the Presumption of Authenticity
A4.2.2.2.1
Evidence for the Presumption of Authenticity
Measure Evidence Against Requirements For Authentic Records
A4.2.2.2.2
Need for Verification
Verify Authenticity
A4.2.2.2.3
Reports on Operation of Permanent Preservation Activity
Technical Description of Records Proposed for Preservation
Assessments of Authenticity
NODE:
A4.2.2.2.2
TITLE:
Assess Authenticity of Records
NUMBER:

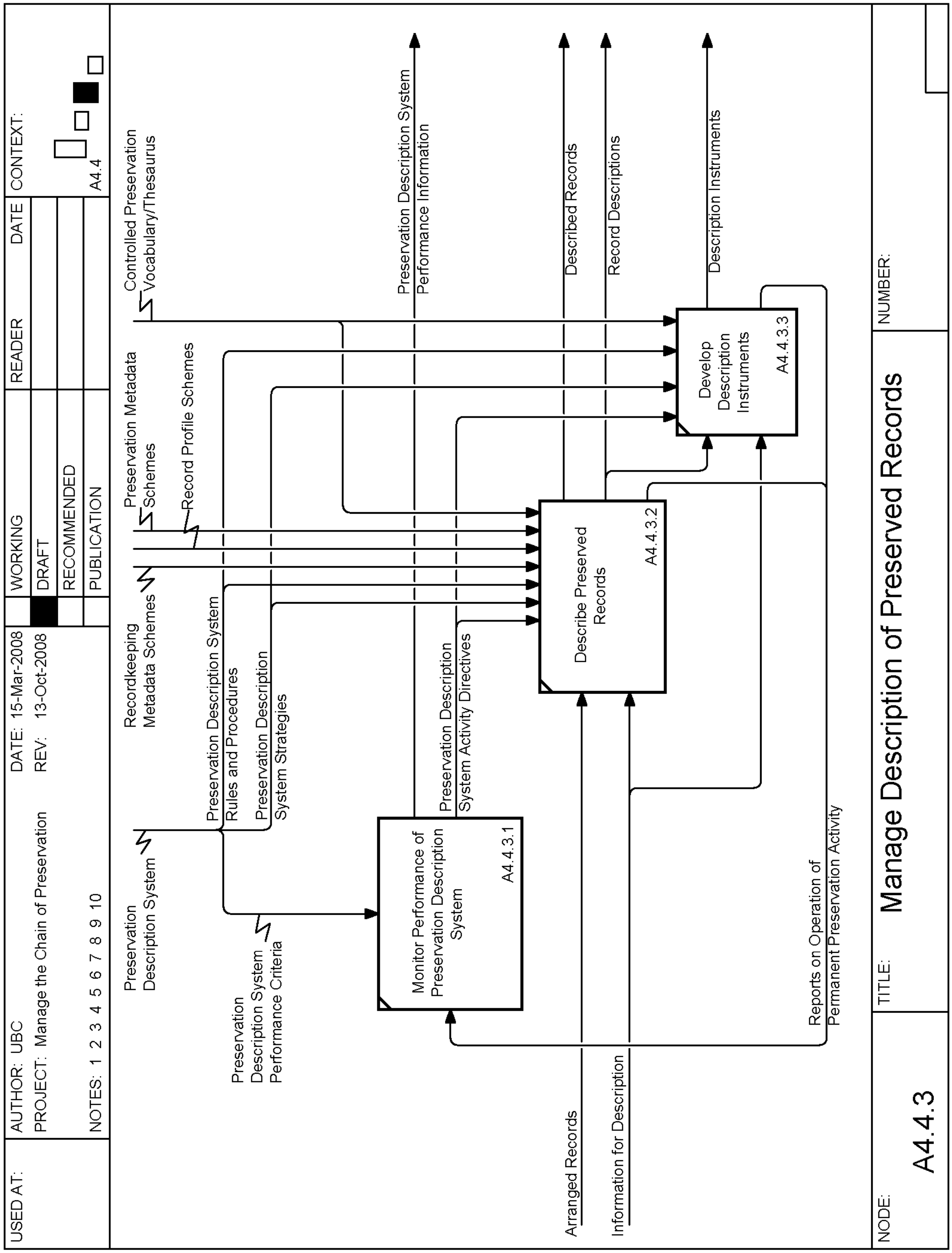

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 28-Mar-2008
REV: 28-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4.2.2
Information for Feasibility
Technical Description of Records Proposed for Preservation
Information About Available Technology
Preservation System Technological Requirements
Reports on Operation of Permanent Preservation Activity
Information About Digital Components to be Preserved
Feasibility Reports
Determine Record Elements to be Preserved
A4.2.2.3.1
Lists of Record Elements to be Preserved
Identify Digital Components to be Preserved
A4.2.2.3.2
Lists of Digital Components to be Preserved
Reconcile Preservation Requirements with Preservation Capabilities
A4.2.2.3.3
NODE:
A4.2.2.3
TITLE:
Determine Feasibility of Preservation
NUMBER:

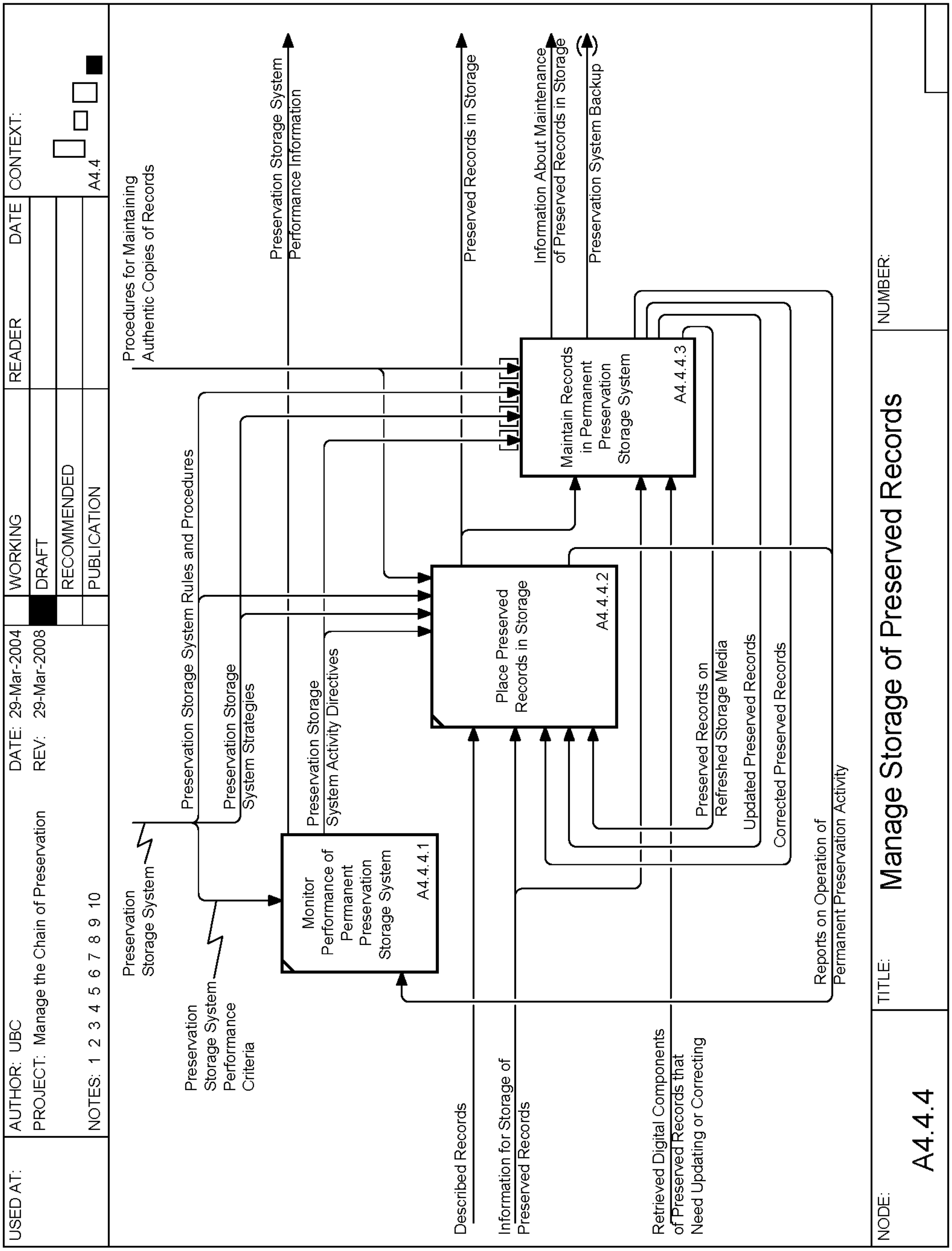

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 13-Feb-2004
REV: 05-Apr-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4
Preservation Acquisition System
Preservation Acquisition System Performance Criteria
Feasibility Reports
Terms and Conditions of Transfer
Preservation Acquisition System Strategies
Preservation Acquisition System Rules and Procedures
Preservation Acquisition System Activity Directives
Assessments of Authenticity
Records Transfer Registration Scheme
Preservation System Access Privileges
Preservation System Instruments
Accession Registration Scheme
Preservation Acquisition System Performance Information
Information for Preservation
Notifications of Rejection of Transfer
Accessioned Records
Information About Accessioned Records
Monitor Performance of Preservation Acquisition System
A4.3.1
Process Records Transfers
A4.3.2
Accession Records
A4.3.3
Records to be Accessioned
Kept Records Transfers
Transfer Documentation for Preservation
Information About Transferred Kept Records
Information About Preserver's Existing Holdings
Reports on Operation of Permanent Preservation Activity
NODE:
A4.3
TITLE:
Acquire Selected Records
NUMBER:

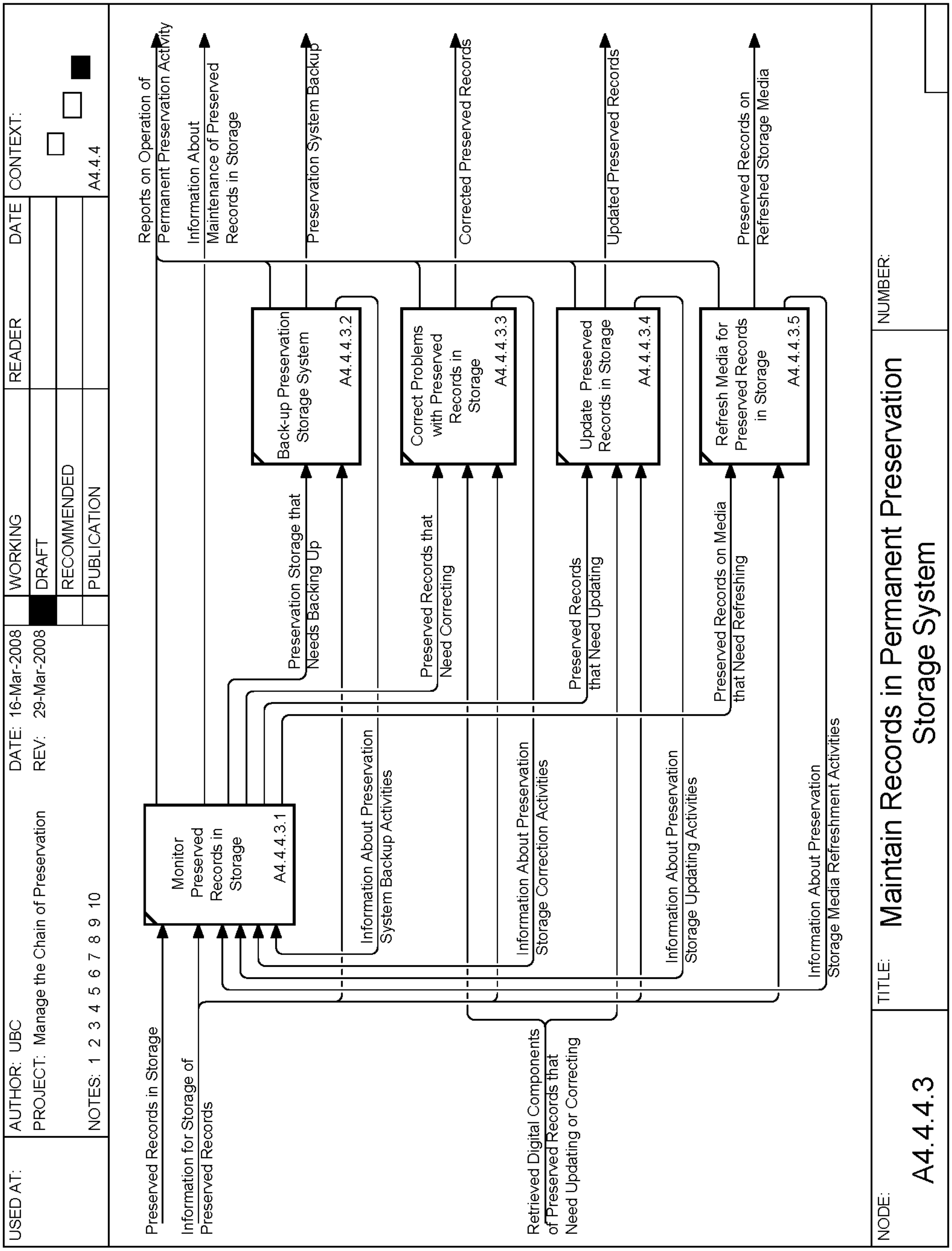

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
DATE: 14-Mar-2008
REV: 05-Apr-2008
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4.3
Records Transfer Registration Scheme
Terms and Conditions of Transfer
Assessments of Authenticity
Feasibility Reports
Reports on Operation of Permanent Preservation Activity
Notifications of Receipt of Transfer
Notifications of Rejection of Transfer
Records to be Accessioned
Information for Preservation
Kept Records Transfers
Transfer Documentation for Preservation
Information About Transferred Kept Records
Register Transfers
A4.3.2.1
Registered Transfers
Confirm Authorization for Transfers
A4.3.2.2
Authorized Transfers
Verify Content of Transfers
A4.3.2.3
Verified Transfers
Confirm Authenticity of Records
A4.3.2.4
Authentic Records
Confirm Feasibility of Preservation
A4.3.2.5
NODE:
A4.3.2
TITLE:
Process Records Transfers
NUMBER:

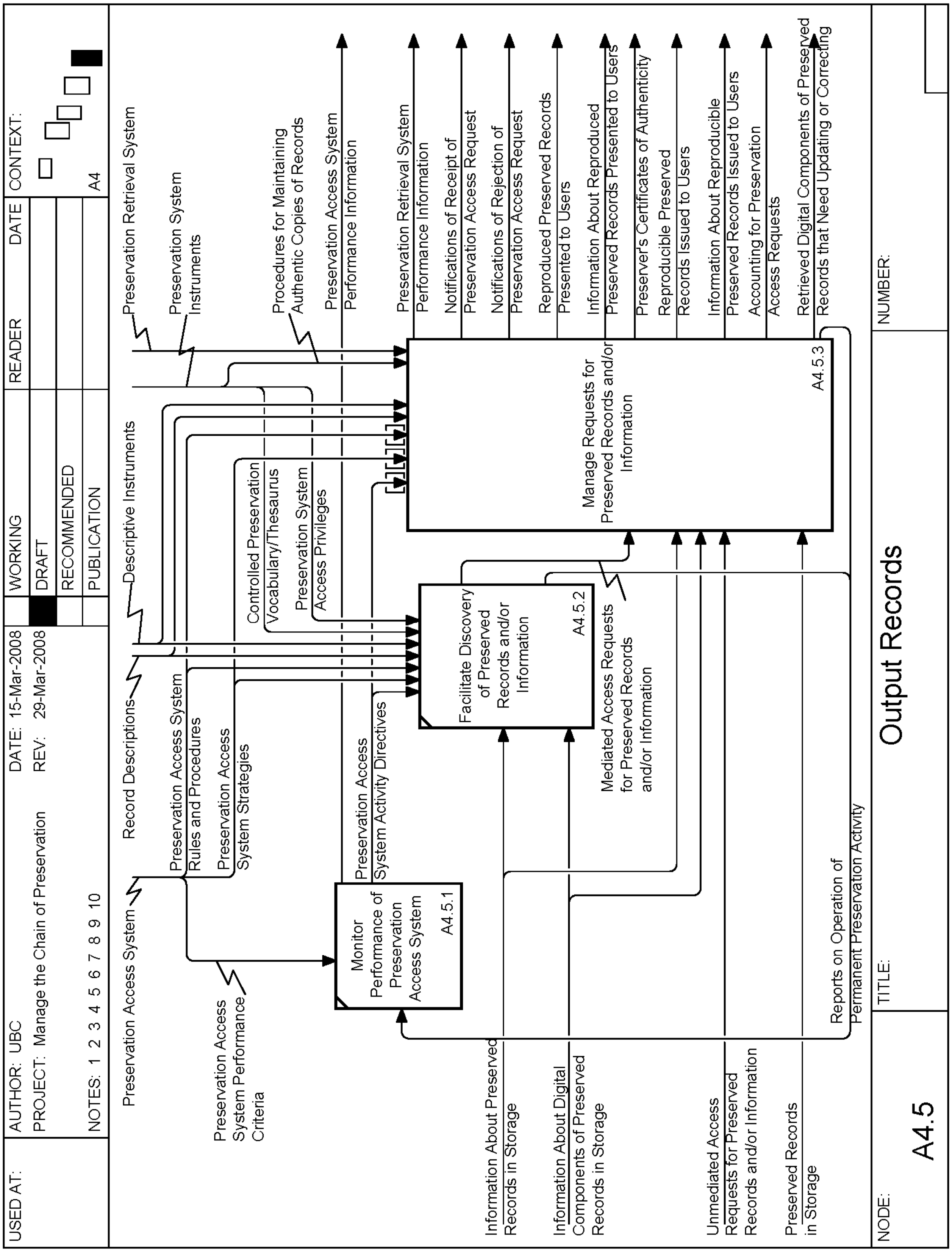

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 13-Feb-2004
REV: 03-Apr-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4
Preservation Information System
Preservation System Instruments
Procedures for Maintaining Authentic Copies of Records
Preservation System Access Privileges
Preservation Metadata Schemes
Controlled Preservation Vocabulary/Thesaurus
Recordkeeping Metadata Schemes
Record Profile Schemes
Preservation Description System
Preservation Storage System
Documentation About Destroyed Records
Information About Transferred Kept Records
Information About Accessioned Records
Accounting for Preservation Access Requests
Notifications of Receipt of Preservation Access Request
Notifications of Rejection of Preservation Access Request
Information for Preservation
Accessioned Records
Retrieved Digital Components of Preserved Records that Need Updating or Correcting
Manage Information About Preserved Records
A4.4.1
Information for Arrangement
Manage Arrangement of Preserved Records
A4.4.2
Arranged Records
Information for Description
Information for Storage of Preserved Records
Information About Maintenance of Preserved Records in Storage
Manage Description of Preserved Records
A4.4.3
Described Records
Manage Storage of Preserved Records
A4.4.4
Preservation Information System Performance Information
Information About Preserved Records in Storage
Information About Digital Components of Preserved Records in Storage
Information About Preserver's Existing Holdings
Preservation Description System Performance Information
Descriptive Instruments
Record Descriptions
Preservation Storage System Performance Information
Preserved Records in Storage
NODE:
A4.4
TITLE:
Preserve Accessioned Records
NUMBER:

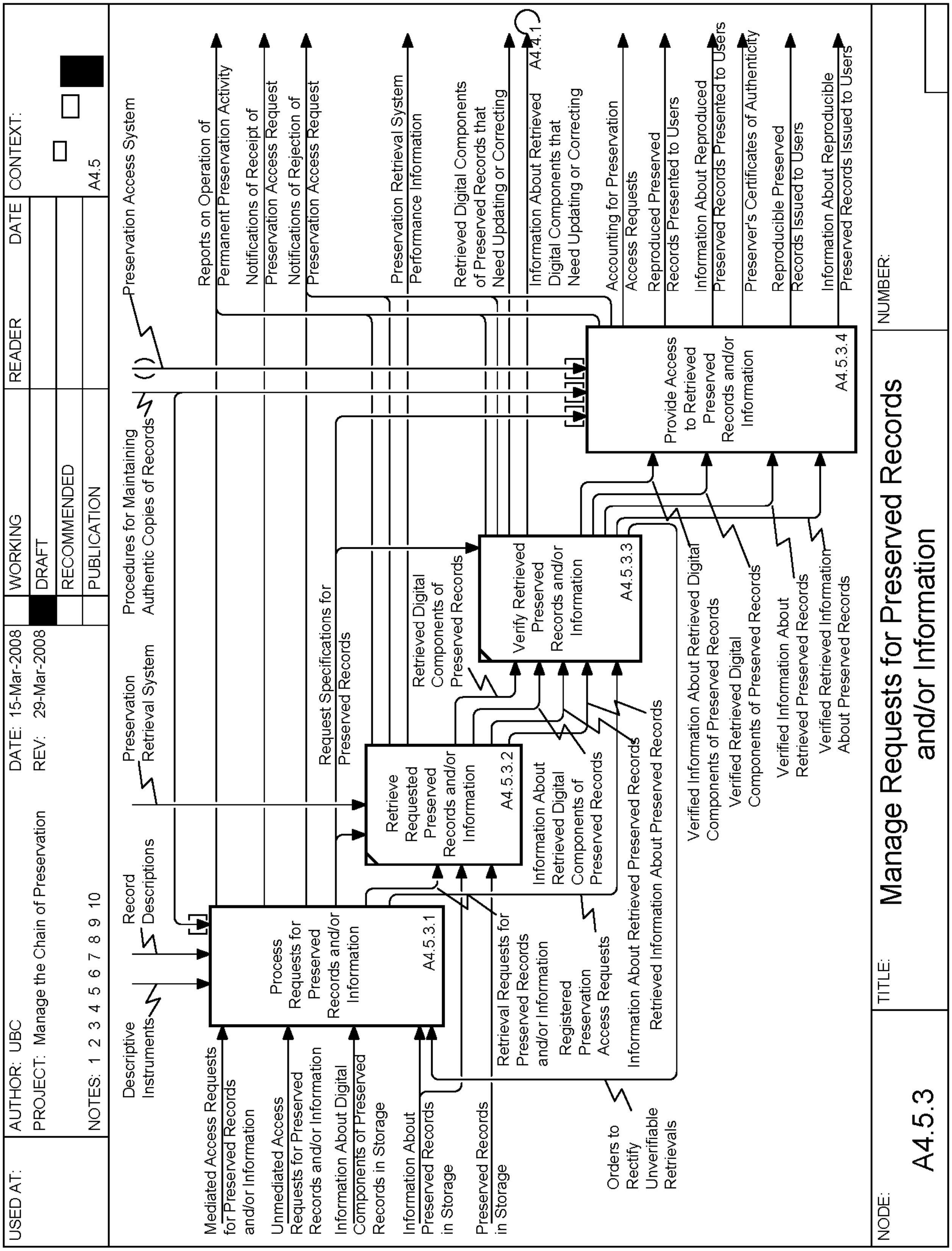

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 29-Mar-2008
REV: 03-Apr-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4.4
Information for Preservation
Documentation About Destroyed Records
Information About Transferred Kept Records
Information About Accessioned Records
Accounting for Preservation Access Requests
Notifications of Rejection of Preservation Access Request
Notifications of Receipt of Preservation Access Request
Information About Retrieved Digital Components that Need Updating or Correcting
A4.5.3
Information About Maintenance of Preserved Records in Storage
Monitor Performance of Preservation Information System
A4.4.1.1
Preservation Information System Activity Directives
Compile Information for Preservation
A4.4.1.2
Update Information on Preservation Actions
A4.4.1.3
Updated Information for Preservation
Reports on Operation of Permanent Preservation Activity
Preservation Information System Performance Information
Information About Preserved Records in Storage
Information About Digital Components of Preserved Records in Storage
Information About Preserver's Existing Holdings
Information for Arrangement
Information for Description
Information for Storage of Preserved Records
NODE:
A4.4.1
TITLE:
Manage Information About Preserved Records
NUMBER:

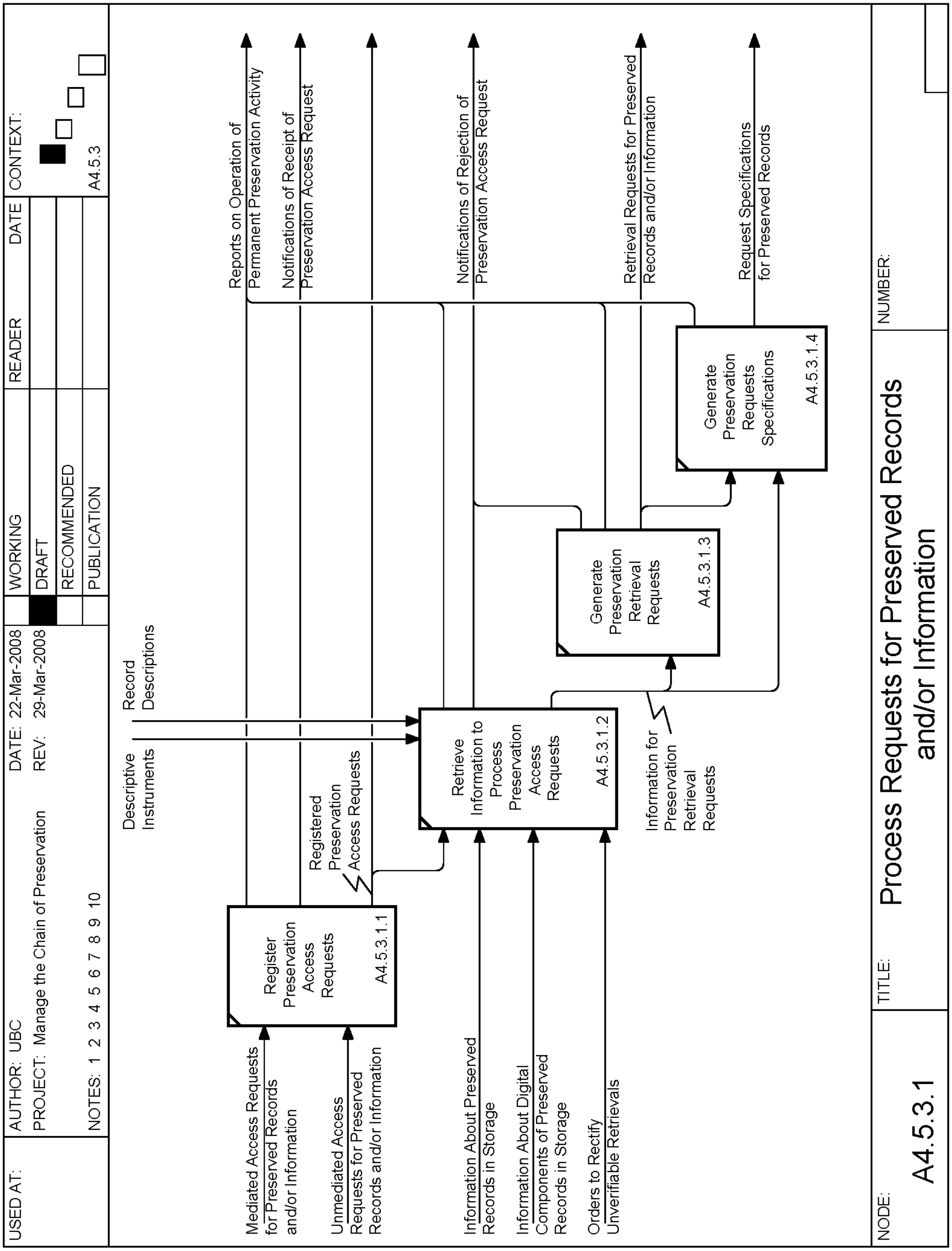

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 15-Mar-2008
REV: 13-Oct-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4.4
Preservation Description System
Preservation Description System Performance Criteria
Recordkeeping Metadata Schemes
Preservation Description System Rules and Procedures
Preservation Description System Strategies
Preservation Metadata Schemes
Record Profile Schemes
Controlled Preservation Vocabulary/Thesaurus
Preservation Description System Activity Directives
Preservation Description System Performance Information
Described Records
Record Descriptions
Description Instruments
Monitor Performance of Preservation Description System
A4.4.3.1
Describe Preserved Records
A4.4.3.2
Develop Description Instruments
A4.4.3.3
Arranged Records
Information for Description
Reports on Operation of Permanent Preservation Activity
NODE:
A4.4.3
TITLE:
Manage Description of Preserved Records
NUMBER:


USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
DATE: 29-Mar-2004
REV: 29-Mar-2008
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4.4
Procedures for Maintaining Authentic Copies of Records
Preservation Storage System Performance Information
Preserved Records in Storage
Information About Maintenance of Preserved Records in Storage
Preservation System Backup
Preservation Storage System Rules and Procedures
Preservation Storage System Strategies
Preservation Storage System Activity Directives
Preservation Storage System
Preservation Storage System Performance Criteria
Monitor Performance of Permanent Preservation Storage System
A4.4.4.1
Place Preserved Records in Storage
A4.4.4.2
Maintain Records in Permanent Preservation Storage System
A4.4.4.3
Described Records
Information for Storage of Preserved Records
Retrieved Digital Components of Preserved Records that Need Updating or Correcting
Preserved Records on Refreshed Storage Media
Updated Preserved Records
Corrected Preserved Records
Reports on Operation of Permanent Preservation Activity
NODE:
A4.4.4
TITLE:
Manage Storage of Preserved Records
NUMBER:

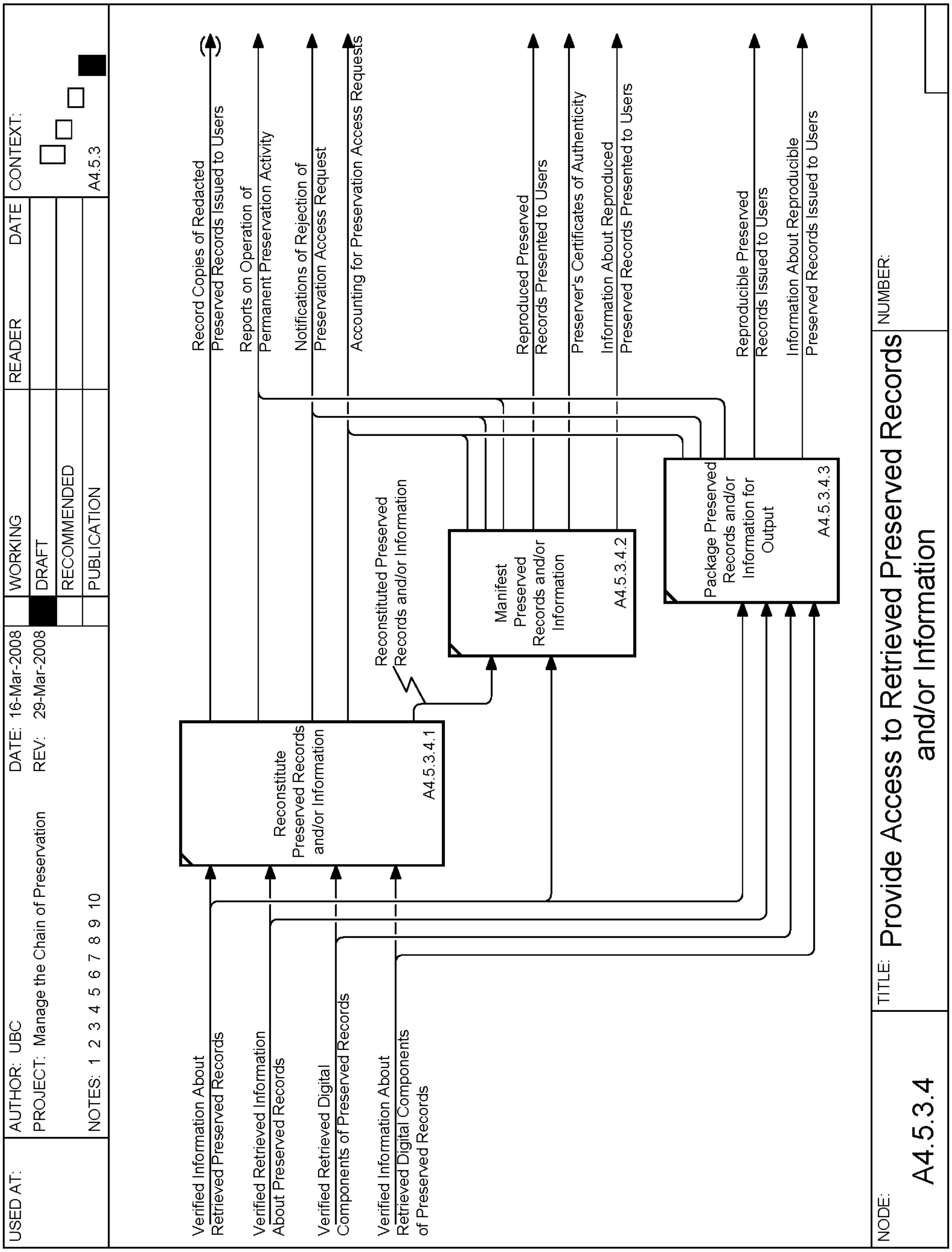

USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
DATE: 16-Mar-2008
REV: 29-Mar-2008
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4.4.4
Preserved Records in Storage
Information for Storage of Preserved Records
Retrieved Digital Components of Preserved Records that Need Updating or Correcting
Monitor Preserved Records in Storage
A4.4.4.3.1
Back-up Preservation Storage System
A4.4.4.3.2
Correct Problems with Preserved Records in Storage
A4.4.4.3.3
Update Preserved Records in Storage
A4.4.4.3.4
Refresh Media for Preserved Records in Storage
A4.4.4.3.5
Preservation Storage that Needs Backing Up
Preserved Records that Need Correcting
Preserved Records that Need Updating
Preserved Records on Media that Need Refreshing
Information About Preservation System Backup Activities
Information About Preservation Storage Correction Activities
Information About Preservation Storage Updating Activities
Information About Preservation Storage Media Refreshment Activities
Reports on Operation of Permanent Preservation Activity
Information About Maintenance of Preserved Records in Storage
Preservation System Backup
Corrected Preserved Records
Updated Preserved Records
Preserved Records on Refreshed Storage Media
NODE:
A4.4.4.3
TITLE:
Maintain Records in Permanent Preservation Storage System
NUMBER:


USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
DATE: 15-Mar-2008
REV: 29-Mar-2008
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4
Preservation Access System
Preservation Access System Performance Criteria
Record Descriptions
Descriptive Instruments
Preservation Access System Rules and Procedures
Preservation Access System Strategies
Controlled Preservation Vocabulary/Thesaurus
Preservation System Access Privileges
Preservation Access System Activity Directives
Preservation Retrieval System
Preservation System Instruments
Procedures for Maintaining Authentic Copies of Records
Preservation Access System Performance Information
Monitor Performance of Preservation Access System
A4.5.1
Facilitate Discovery of Preserved Records and/or Information
A4.5.2
Manage Requests for Preserved Records and/or Information
A4.5.3
Mediated Access Requests for Preserved Records and/or Information
Information About Preserved Records in Storage
Information About Digital Components of Preserved Records in Storage
Unmediated Access Requests for Preserved Records and/or Information
Preserved Records in Storage
Reports on Operation of Permanent Preservation Activity
Preservation Retrieval System Performance Information
Notifications of Receipt of Preservation Access Request
Notifications of Rejection of Preservation Access Request
Reproduced Preserved Records Presented to Users
Information About Reproduced Preserved Records Presented to Users
Preserver's Certificates of Authenticity
Reproducible Preserved Records Issued to Users
Information About Reproducible Preserved Records Issued to Users
Accounting for Preservation Access Requests
Retrieved Digital Components of Preserved Records that Need Updating or Correcting
NODE:
A4.5
TITLE:
Output Records
NUMBER:


USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 15-Mar-2008
REV: 29-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4.5
Mediated Access Requests for Preserved Records and/or Information
Unmediated Access Requests for Preserved Records and/or Information
Information About Digital Components of Preserved Records in Storage
Information About Preserved Records in Storage
Preserved Records in Storage
Orders to Rectify Unverifiable Retrievals
Descriptive Instruments
Record Descriptions
Preservation Retrieval System
Request Specifications for Preserved Records
Procedures for Maintaining Authentic Copies of Records
Preservation Access System
Process Requests for Preserved Records and/or Information
A4.5.3.1
Retrieve Requested Preserved Records and/or Information
A4.5.3.2
Verify Retrieved Preserved Records and/or Information
A4.5.3.3
Provide Access to Retrieved Preserved Records and/or Information
A4.5.3.4
Retrieval Requests for Preserved Records and/or Information
Registered Preservation Access Requests
Information About Retrieved Preserved Records
Retrieved Information About Preserved Records
Information About Retrieved Digital Components of Preserved Records
Retrieved Digital Components of Preserved Records
Verified Information About Retrieved Digital Components of Preserved Records
Verified Retrieved Digital Components of Preserved Records
Verified Information About Retrieved Preserved Records
Verified Retrieved Information About Preserved Records
Reports on Operation of Permanent Preservation Activity
Notifications of Receipt of Preservation Access Request
Notifications of Rejection of Preservation Access Request
Preservation Retrieval System Performance Information
Retrieved Digital Components of Preserved Records that Need Updating or Correcting
Information About Retrieved Digital Components that Need Updating or Correcting
A4.4.1
Accounting for Preservation Access Requests
Reproduced Preserved Records Presented to Users
Information About Reproduced Preserved Records Presented to Users
Preserver's Certificates of Authenticity
Reproducible Preserved Records Issued to Users
Information About Reproducible Preserved Records Issued to Users
NODE:
A4.5.3
TITLE:
Manage Requests for Preserved Records and/or Information
NUMBER:



USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
DATE: 22-Mar-2008
REV: 29-Mar-2008
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4.5.3
Mediated Access Requests for Preserved Records and/or Information
Unmediated Access Requests for Preserved Records and/or Information
Information About Preserved Records in Storage
Information About Digital Components of Preserved Records in Storage
Orders to Rectify Unverifiable Retrievals
Descriptive Instruments
Record Descriptions
Register Preservation Access Requests
A4.5.3.1.1
Registered Preservation Access Requests
Retrieve Information to Process Preservation Access Requests
A4.5.3.1.2
Information for Preservation Retrieval Requests
Generate Preservation Retrieval Requests
A4.5.3.1.3
Generate Preservation Requests Specifications
A4.5.3.1.4
Reports on Operation of Permanent Preservation Activity
Notifications of Receipt of Preservation Access Request
Notifications of Rejection of Preservation Access Request
Retrieval Requests for Preserved Records and/or Information
Request Specifications for Preserved Records
NODE:
A4.5.3.1
TITLE:
Process Requests for Preserved Records and/or Information
NUMBER:


USED AT:
AUTHOR: UBC
PROJECT: Manage the Chain of Preservation
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 16-Mar-2008
REV: 29-Mar-2008
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A4.5.3
Verified Information About Retrieved Preserved Records
Verified Retrieved Information About Preserved Records
Verified Retrieved Digital Components of Preserved Records
Verified Information About Retrieved Digital Components of Preserved Records
Reconstitute Preserved Records and/or Information
A4.5.3.4.1
Reconstituted Preserved Records and/or Information
Manifest Preserved Records and/or Information
A4.5.3.4.2
Package Preserved Records and/or Information for Output
A4.5.3.4.3
(A)
Record Copies of Redacted Preserved Records Issued to Users
Reports on Operation of Permanent Preservation Activity
Notifications of Rejection of Preservation Access Request
Accounting for Preservation Access Requests
Reproduced Preserved Records Presented to Users
Preserver's Certificates of Authenticity
Information About Reproduced Preserved Records Presented to Users
Reproducible Preserved Records Issued to Users
Information About Reproducible Preserved Records Issued to Users
NODE:
A4.5.3.4
TITLE:
Provide Access to Retrieved Preserved Records and/or Information
NUMBER:

Chain of Preservation Model Activity Definitions

A0, Manage Chain of Preservation
To design, implement and maintain a framework to control the records throughout the processes of creation, maintenance and use, disposition and preservation.

A1, Manage Framework for Chain of Preservation
To determine framework requirements, and design, implement and maintain a chain of preservation framework.

A1.1, Develop Management Framework
To analyze information about the records creator and its existing records and about the designated records preserver to identify the policies and requirements for the chain of preservation framework.

A1.1.1, Analyze Records Creator
To assess the information concerning the records creator's mission, organizational structure, activities and existing technological, financial and human resources, and records related needs and risks to help identify the requirements for the chain of preservation framework.

A1.1.2, Analyze Creator's Existing Records
To assess information about creator's existing records to determine framework requirements.

A1.1.3, Analyze Designated Records Preserver
To assess the information concerning the designated records preserver's mission, organizational structure, activities and existing technological, financial and human resources, and records preservation related needs and risks to help identify the requirements for the chain of preservation framework.

A1.1.4, Establish Management Policies
To develop and document management regime policies for establishing overall framework design requirements.

A1.1.5, Establish Design Requirements
To identify the rules guiding the chain of preservation framework on the basis of the analysis of the records creator and its existing records.

A1.2, Design Framework
To develop a record-making system design, recordkeeping system design and permanent preservation system design.

A1.2.1, Design Record-making System
To develop the record-making system's administrative infrastructure, determine functional requirements for the system, establish performance criteria for the system and develop the functional infrastructure design for the system.

A1.2.1.1, Develop Record-making System Administrative Infrastructure
To define, analyze, create and document a comprehensive, integrated set of administrative policies, strategies, rules and procedures, and instruments to support record-making activities and to enable the record-making system to meet its functional requirements.

A1.2.1.1.1, Develop Record-making System Policies
To determine and document the collective, high-level management principles that guide and control development, implementation and execution of record-making system activities.

A1.2.1.1.2, Develop Record-making System Strategies

To determine and document the authoritative objectives and methods governing the operation of the record-making system.

A1.2.1.1.3, Develop Record-making System Rules and Procedures

To determine and document the authoritative instructions governing the operation of the record-making system.

A1.2.1.1.4, Develop Record-making System Instruments

To define, analyze, create and document the various administrative tools that support record-making processes, such as metadata schemes and records forms.

A1.2.1.2, Establish Record-making System Functional Requirements

To develop and document comprehensive and integrated performance, monitoring and technological requirements for the record-making system.

A1.2.1.2.1, Determine Record-making System Performance Requirements

To identify the operational and administrative specifications for measuring the continuing ability of the record-making system to fulfil its purpose.

A1.2.1.2.2, Determine Record-making System Monitoring Requirements

To identify the operational and administrative conditions for providing ongoing assessment of the operation of the record-making system in relation to the established system performance requirements.

A1.2.1.2.3, Determine Record-making System Technological Requirements

To specify the hardware and software needed for the record-making system.

A1.2.1.3, Establish Record-making System Performance Criteria

To develop operational benchmarks or standards for operation of the record-making system, in relation to established requirements, against which the continuing performance and adequacy of an activity, function, process, sub-system or structure within the system can be measured.

A1.2.1.4, Design Record-making System Functional Infrastructure

To develop a comprehensive, integrated design for the record-making system and each of its documents and records capture, identification, declaration, execution and transfer sub-systems.

A1.2.2, Design Recordkeeping System

To develop the recordkeeping system's administrative infrastructure, determine functional requirements for the system, establish performance criteria for the system and develop the functional infrastructure design for the system.

A1.2.2.1, Develop Recordkeeping System Administrative Infrastructure

To define, analyze, create and document a comprehensive, integrated set of administrative policies, strategies, rules and procedures, and instruments to support recordkeeping activities and to enable the recordkeeping system to meet its functional requirements.

A1.2.2.1.1, Develop Recordkeeping System Policies

To determine and document the collective, high-level management principles that guide and control development, implementation and execution of recordkeeping system activities.

A1.2.2.1.2, Develop Recordkeeping System Strategies

To determine and document the authoritative objectives and methods governing the operation of the record-keeping system.

A1.2.2.1.3, Develop Recordkeeping System Rules and Procedures

To determine and document the authoritative instructions governing the operation of the recordkeeping system.

A1.2.2.1.4, Develop Recordkeeping System Instruments

To define, analyze, create and document the various administrative tools that support recordkeeping processes, such as metadata schemes, registration and classification schemes, a retention schedule and a controlled vocabulary.

A1.2.2.1.4.1, Develop Recordkeeping Schemes

To establish the metadata, registration and classification schemes used in the recordkeeping system.

A1.2.2.1.4.2, Develop Retention Schedule

To determine and document the disposition of each series and/or class of records.

A1.2.2.1.4.3, Develop Recordkeeping Indexes

To define, analyze, create and document the tools, such as a controlled vocabularies and thesauri, that facilitate efficient and effective location of information, records and/or records aggregates in the recordkeeping system suited to a particular inquiry or purpose.

A1.2.2.2, Establish Recordkeeping System Functional Requirements

To develop and document comprehensive and integrated performance, monitoring and technological requirements for the recordkeeping system.

A1.2.2.2.1, Determine Recordkeeping System Performance Requirements

To identify the operational and administrative specifications for measuring the continuing ability of the recordkeeping system to fulfil its purpose.

A1.2.2.2.2, Determine Recordkeeping System Monitoring Requirements

To identify the operational and administrative conditions for providing ongoing assessment of the operation of the recordkeeping system in relation to the established system performance requirements.

A1.2.2.2.3, Determine Recordkeeping System Technological Requirements

To specify the hardware and software needed for the recordkeeping system.

A1.2.2.3, Establish Recordkeeping System Performance Criteria

To develop operational benchmarks or standards for operation of the record-making system, in relation to established system performance, monitoring and technological requirements, against which the continuing performance and adequacy of an activity, function, process, sub-system or structure within the system can be measured.

A1.2.2.4, Design Recordkeeping System Functional Infrastructure

To develop a comprehensive, integrated design for the recordkeeping system and each of its records information, storage, retrieval, access and disposition sub-systems.

A1.2.3, Design Permanent Preservation System

To develop the permanent preservation system's administrative infrastructure, determine functional requirements for the system, establish performance criteria for the system and develop the functional infrastructure design for the system.

A1.2.3.1, Develop Preservation System Administrative Infrastructure

To define, analyze, create and document a comprehensive, integrated set of administrative policies, strategies, rules and procedures, and instruments to support preservation activities and to enable the permanent preservation system to meet its functional requirements.

A1.2.3.1.1, Develop Preservation System Policies
To determine and document the collective, high-level management principles that guide and control development, implementation and execution of permanent preservation system activities.

A1.2.3.1.2, Develop Preservation System Strategies
To determine and document the authoritative objectives and methods governing the operation of the permanent preservation system.

A1.2.3.1.3, Develop Preservation System Rules and Procedures
To determine and document the authoritative instructions governing the operation of the permanent preservation system.

A1.2.3.1.4, Develop Preservation System Instruments
To define, analyze, create and document the various administrative tools that support preservation processes, such as metadata schemes, transfer and accession registration schemes and a controlled vocabulary.

A1.2.3.2, Establish Preservation System Functional Requirements
To identify and document comprehensive and integrated performance, monitoring and technological requirements for the permanent preservation system.

A1.2.3.2.1, Determine Preservation System Performance Requirements
To identify the operational and administrative specifications for measuring the continuing ability of the permanent preservation system to fulfil its purpose.

A1.2.3.2.2, Determine Preservation System Monitoring Requirements
To identify the needs for providing ongoing assessment of the operation of the permanent preservation system in relation to the operational and administrative procedures and instruments developed for meeting these needs.

A1.2.3.2.3, Determine Preservation System Technological Requirements
To specify the hardware and software needed for the permanent preservation system.

A1.2.3.3, Establish Preservation System Performance Criteria
To develop operational benchmarks or standards for operation of the permanent preservation system, in relation to established system performance, monitoring and technological requirements, against which the continuing performance and adequacy of an activity, function, process, sub-system or structure within the system can be measured.

A1.2.3.4, Design Preservation System Functional Infrastructure
To develop a comprehensive, integrated design for the permanent preservation system and each of its records information, selection, acquisition, description, storage, retrieval and access sub-systems.

A1.3, Implement Framework
To acquire, test and activate all the components of the record-making, recordkeeping, and permanent preservation systems, and issue information about implementation problems.

A1.4, Maintain Framework
To assess information about the performance of the record-making, recordkeeping and permanent preservation systems and to make recommendations on the revision of the overall framework design and/or its constituent systems.

A2, Manage Records in a Record-making System
To provide overall control and co-ordination of activities in the record-making system, including the creation and setting aside of records, and monitoring of the performance of the record-making system.

A2.1, Monitor Performance of Record-making System

To assess the efficacy of the performance of the record-making system by analyzing performance reports on the operation of each of the record-making system's sub-systems and issue activity directives for record-making activities and information on the performance of the record-making system for use in continued maintenance of the chain of preservation framework.

A2.2, Manage Making and Receipt of Records

To provide overall control and co-ordination of document and record making and receipt activities, including the capture and identification of documents made or received by the creator and their subsequent declaration and execution as records.

A2.2.1, Make Documents

To compile digital information in a syntactic manner in accordance with the specifications of the creator's documentary forms, integrated business and documentary procedures and record-making access privileges.

A2.2.2, Capture Documents

To record and save (i.e., affix to a digital medium in a stable syntactic manner) particular instantiations of incoming external documents or internal documents made by the creator in the record-making system in accordance with the specifications of the creator's integrated business and documentary procedures and record-making access privileges.

A2.2.3, Identify Documents

To attach to each document identity metadata that convey the action in which the document participates and its immediate context.

A2.2.4, Declare Records

To intellectually set aside records by assigning classification codes from the classification scheme to made or received documents and adding these codes to the identifying metadata and by assigning to the documents registration numbers based on the registration scheme, and adding these numbers to the identifying metadata.

A2.2.5, Execute Records

To attach to each record metadata that convey information related to, and actions taken during the course of, the formal execution phase of the administrative procedure in which the record participates, which may also involve transmitting documents to external physical or juridical persons and making record copies of the sent documents.

A2.3, Manage Setting Aside of Completed Records

To provide overall control and co-ordination of the transfer of executed or completed records to the recordkeeping system by preparing the records for transfer, transferring the records and monitoring the performance of the record-making transfer system.

A2.3.1, Monitor Performance of Record-making Transfer System

To assess the efficacy of the performance of the record-making transfer system by analyzing reports on the operation of record-making activities, and issue activity directives for transfer activities and issue information on the performance of the record-making transfer system for use in continued maintenance of the record-making system.

A2.3.2, Prepare Completed Records for Transfer to Recordkeeping System

To attach to completed records integrity and related metadata that convey information related to, and actions taken during the course of, managing the records for records management purposes prior to setting them aside in the recordkeeping system; compile information about the records that is needed to meet all transfer information requirements; and ensure that the records are in the proper format for transfer to the recordkeeping system as prescribed by recordkeeping system rules and procedures and technological requirements.

A2.3.3, Transfer Completed Records to Recordkeeping System

To send or transmit completed records prepared for transfer to the office responsible for the recordkeeping function with the accompanying documentation necessary for recordkeeping.

A3, Manage Records in a Recordkeeping System

To provide overall control and co-ordination of activities in the recordkeeping system, including records storage, retrieval and access, disposition, and monitoring of the performance of the recordkeeping system.

A3.1, Monitor Performance of Recordkeeping System

To assess the efficacy of the performance of the recordkeeping system by analyzing performance reports on the operation of recordkeeping sub-system activities, and issue activity directives for recordkeeping activities and information on the performance of the recordkeeping system for use in continued maintenance of the chain of preservation framework.

A3.2, Manage Maintenance of Kept Records

To provide overall control and co-ordination of the recordkeeping storage system and the records stored in the system by managing information about kept records and their digital components, placing the records in storage, maintaining the digital components and monitoring the performance of the storage system.

A3.2.1, Manage Information About Kept Records

To compile information about records in the recordkeeping system and about records maintenance activities and to provide overall control and co-ordination of that information for use in records appraisal activities by the preserver and in records indexing, storage, access and disposition activities by the creator.

A3.2.2, Manage Indexing of Kept Records

To provide overall control and co-ordination of records indexing activities, including monitoring the indexing system, indexing kept records and developing indexing instruments to help facilitate records discovery and retrieval.

A3.2.2.1, Monitor Performance of Recordkeeping Indexing System

To assess the efficacy of the performance of the recordkeeping indexing system by analyzing reports on the operation of recordkeeping activities, and issue activity directives for indexing activities and information on the performance of the indexing system for use in continued maintenance of the recordkeeping system.

A3.2.2.2, Index Kept Records

To establish and record access points for kept records within the context of a controlled recordkeeping vocabulary applied according to recordkeeping indexing system rules, procedures and strategies.

A3.2.2.3, Develop Indexing Instruments

To prepare tools that facilitate discovery and retrieval of the records in the recordkeeping system, such as guides, inventories and indexes.

A3.2.3, Manage Storage of Kept Records

To provide overall control and co-ordination of the recordkeeping storage system and the records stored in the system by placing the records in storage, maintaining their digital components and monitoring the performance of the storage system.

A3.2.3.1, Monitor Performance of Recordkeeping Storage System

To assess the efficacy of the performance of the recordkeeping storage system by analyzing reports on the operation of recordkeeping activities, and issue activity directives for storage activities and information on the performance of the recordkeeping storage system for use in continued maintenance of the recordkeeping system.

A3.2.3.2, Place Kept Records in Storage

To place the digital components of kept records and their metadata into storage in accordance with the procedures for maintaining authentic records and the actions prescribed by the recordkeeping storage system strategies, rules and procedures and activity directives.

A3.2.3.3, Maintain Records in Recordkeeping Storage System

To monitor the storage of kept records and their digital components and metadata, periodically back-up the recordkeeping storage system and, as necessary, correct problems with and update the digital components, and/or refresh storage media to ensure the records in the system remain accessible, legible and intelligible over time.

A3.2.3.3.1, Monitor Kept Records in Storage

To keep track of the condition and maintenance requirements of kept records and their digital components--more specifically, their digital components and metadata--and the media on which they are stored in the recordkeeping storage system to identify storage that needs backing-up, digital components and/or metadata that need correcting or updating and media that need refreshing; and to issue reports on maintenance activities.

A3.2.3.3.2, Back-up Recordkeeping Storage System

To routinely make a copy of all digital content in the recordkeeping storage system, including the operating system, the software applications and all digital objects in the system, for the purpose of recovery in the event of a disaster resulting in system failure or corruption, and record information about these back-up activities.

A3.2.3.3.3, Correct Problems with Kept Records in Storage

To take the actions prescribed by the recordkeeping storage system strategies, rules and procedures and activity directives, in accordance with the procedures for maintaining authentic records, to eliminate problems in storage, and record information about these correction activities.

A3.2.3.3.4, Update Kept Records in Storage

To carry out conversion actions on the digital components of stored kept records in accordance with the procedures for maintaining authentic records and the actions prescribed by the recordkeeping storage system strategies, rules and procedures and activity directives, to ensure the records remain accessible, legible and intelligible over time (such as by migration, standardization or transformation to persistent form), and record information about these updating activities.

A3.2.3.3.5, Refresh Media for Kept Records in Storage

To copy or transfer the digital components of kept records in storage from one medium to another, or otherwise ensure the storage medium remains sound, in accordance with the procedures for maintaining authentic records and the actions prescribed by the recordkeeping storage system strategies, rules and procedures and activity directives, and record information about these media refreshment activities.

A3.3, Manage Access to Kept Records

To facilitate discovery of, and manage requests for, kept records and/or information about kept records, and monitor the performance of the recordkeeping access system.

A3.3.1, Monitor Performance of Recordkeeping Access System

To assess the efficacy of the performance of the recordkeeping access system by analyzing reports on the operation of recordkeeping activities, and issue activity directives for access activities and information on the performance of the recordkeeping access system for use in continued maintenance of the recordkeeping system.

A3.3.2, Facilitate Discovery of Kept Records and/or Information

To provide authorized internal and external users access to, and assistance in the use of, the tools and resources necessary to support querying and searching for, and discovery of, information, records and/or records aggregates in the recordkeeping system suited to a particular inquiry or purpose.

A3.3.3, Manage Requests for Kept Records and/or Information

To provide overall control and co-ordination of internal and external requests for access to records and/or information about kept records by processing access requests, retrieving digital components for requested records and/or information, verifying retrieved components and information and providing access to retrieved records and/or information.

A3.3.3.1, Process Requests for Kept Records and/or Information

To register access requests for kept records and/or information, translate them, define request specifications, generate retrieval requests and account for any problems with processing requests.

A3.3.3.1.1, Register Recordkeeping Access Requests

To record registration information about received requests for access to kept records and/or information about the records and issue notifications of receipt to the persons requesting the records.

A3.3.3.1.2, Retrieve Information to Process Recordkeeping Access Requests

To gather the information, from indexing instruments, record profiles and other recordkeeping tools, needed to process access requests for kept records and/or information about records.

A3.3.3.1.3, Generate Recordkeeping Retrieval Requests

To translate access requests for kept records and/or information into requests to the recordkeeping storage and information systems for retrieval of the exact digital components and/or information required to fulfil the access requests.

A3.3.3.1.4, Generate Recordkeeping Requests Specifications

To issue instructions to the recordkeeping retrieval and access systems on how to fulfil requests for kept records and/or information about the records based on analyses of the requests and processing information in relation to recordkeeping access system strategies, rules and procedures (including procedures for maintaining authentic records) and access privileges.

A3.3.3.2, Retrieve Requested Kept Records and/or Information

To output copies of digital components of records, information about digital components of records, rendering information about records and/or content information about records retrieved from storage in the recordkeeping system in response to retrieval requests for components and/or information.

A3.3.3.3, Verify Retrieved Kept Records and/or Information

To determine whether all components and information necessary to satisfy requests for kept records and/or information about kept records have been received and can be processed for output and, in cases where digital components are encountered that need updating or correcting, redirect them (or information about the problems encountered) to the maintenance function of the recordkeeping storage system.

A3.3.3.4, Provide Access to Retrieved Kept Records and/or Information

To fulfil access requests by either reconstituting the retrieved digital components of kept records and/or information in authentic form and presenting the manifested records or information to users, or by packaging the retrieved digital components with information about how to reconstitute and present the records and/or information with the appropriate extrinsic form and issuing the packaged materials to users, and account for the success or failure of either activity.

A3.3.3.4.1, Reconstitute Kept Records and/or Information

To link or assemble all the verified digital components of requested kept records and/or information about kept records as necessary to reproduce and present the records and/or information in authentic form and, if necessary, redact records and/or information to meet privacy and/or copyright requirements.

A3.3.3.4.2, Manifest Kept Records and/or Information

To present copies of the reconstituted requested kept records and/or requested information about the records with the appropriate extrinsic form and with information about their relationships to one another (archival bond) and, if requested, produce a Certificate of Authenticity for the records copies.

A3.3.3.4.3, Package Kept Records and/or Information for Output

To combine the digital components of the requested kept records and/or requested information about kept records with information on how to reconstitute and manifest the records or information with the appropriate extrinsic form.

A3.4, Manage Disposition of Kept Records

To provide overall control and co-ordination of records disposition activities, including monitoring the performance of the disposition system, processing disposition information and, in accordance with disposition activity directives and disposition rules and procedures, destroying kept records and/or preparing and transferring kept records to the designated preserver.

A3.4.1, Monitor Performance of Disposition System

To assess the efficacy of the performance of the recordkeeping disposition system by analyzing reports on the operation of recordkeeping activities, and issue activity directives for disposition activities and information on the performance of the recordkeeping storage system for use in continued maintenance of the recordkeeping system.

A3.4.2, Identify Kept Records for Disposition

To identify records and information about records in the recordkeeping system earmarked either for destruction or transfer to the designated preserver, as determined by the creator's retention schedule.

A3.4.3, Destroy Kept Records

To obliterate kept records, and information related to the records, identified for destruction and provide documentation about the records destroyed.

A3.4.4, Prepare Kept Records for Transfer to Designated Preserver

To attach to kept records integrity and related metadata about actions taken during the course of preparing the records for transfer to the designated preserver in accordance with the terms and conditions of transfer, and compile information about the records that is needed to meet all transfer information requirements.

A3.4.5, Transfer Kept Records to Designated Preserver

To send or transmit kept records prepared for transfer to permanent preserver (or, as applicable, the office of the creator responsible for the permanent preservation function) with the accompanying documentation necessary for permanent preservation.

A4, Manage Records in a Permanent Preservation System

To provide overall control and co-ordination of activities in the permanent preservation system, including records appraisal and selection, acquisition, description, storage, retrieval and access, and monitoring of the performance of the permanent preservation system.

A4.1, Monitor Performance of Permanent Preservation System

To assess the efficacy of the performance of the permanent preservation system by analyzing performance reports on the operation of permanent preservation sub-system activities, and issue activity directives for preservation activities and information on the performance of the permanent preservation system for use in continued maintenance of the chain of preservation framework.

A4.2, Appraise Records for Permanent Preservation

To make appraisal decisions by compiling information about kept records and their context, assessing their value, and determining the feasibility of their preservation; and to monitor appraised records and appraisal decisions to identify any necessary changes to appraisal decisions over time.

A4.2.1, Monitor Performance of Preservation Selection System

To assess the efficacy of the performance of the permanent preservation selection system by analyzing reports on the operation of preservation activities, and issue activity directives for selection activities and information on the performance of the permanent preservation selection system for use in continued maintenance of the permanent preservation system.

A4.2.2, Analyze Kept Records for Preservation

To assess information concerning the kept records being appraised, including their contexts, value and preservation feasibility.

A4.2.2.1, Analyze Information About Records

To collect, organise, record and assess relevant information from the kept records being appraised and about their juridical-administrative, provenancial, procedural, documentary and technological contexts.

A4.2.2.2, Assess Value of Records

To analyze and judge: (1) the capacity of records being appraised to serve the continuing interests of their creator and society; and (2) the grounds for presuming the records to be authentic.

A4.2.2.2.1, Assess Continuing Value of Records

To analyze and judge the capacity of records being appraised to serve the continuing interests of their creator and society.

A4.2.2.2.2, Assess Authenticity of Records

To analyze and judge the grounds for presuming records being appraised to be authentic.

A4.2.2.2.2.1, Compile Evidence Supporting the Presumption of Authenticity

To collect, organize and record evidence of the identity and integrity of records being appraised and about the procedural controls applied to them, to support the presumption of authenticity of those records.

A4.2.2.2.2.2, Measure Evidence Against Requirements For Authentic Records
To compare the evidence compiled about the identity, integrity and procedural controls of the records being appraised with the requirements for authentic records.

A4.2.2.2.2.3, Verify Authenticity
To use verification methods to determine the authenticity of records being appraised in cases where there is insufficient evidence to meet the requirements for presuming the authenticity of records.

A4.2.2.2.3, Determine Value of Records
To establish the value of records being appraised based on assessments of their continuing value and their authenticity.

A4.2.2.3, Determine Feasibility of Preservation
To identify the elements and digital components of the records being appraised, reconcile their preservation requirements with the preserver's current and anticipated preservation capabilities, and provide documentation about the digital components to be preserved and the feasibility of preservation.

A4.2.2.3.1, Determine Record Elements to be Preserved
To identify the necessary documentary components (e.g., record profile, attachments, annotations, etc.) and elements of form (e.g., author, date, subject line, etc.) of records to be preserved to determine which record elements must be preserved to protect the authenticity of those records.

A4.2.2.3.2, Identify Digital Components to be Preserved
To identify the digital components that manifest the record elements that need to be preserved to protect the authenticity of records selected for permanent preservation.

A4.2.2.3.3, Reconcile Preservation Requirements with Preservation Capabilities
To determine whether the digital components manifesting the record elements that need to be preserved to protect the authenticity of records selected for permanent preservation can in fact be preserved given the preserver's current and anticipated preservation capabilities.

A4.2.3, Make Appraisal Decisions
To decide on and document the retention and disposition of records based on valuation and feasibility information, and to agree on and document the terms and conditions of transfer of the records to the preserver.

A4.2.4, Monitor Appraisal Decisions
To keep track of appraisal decisions in relation to subsequent developments within the creator's and/or preserver's activities that might make it necessary to adjust or redo an appraisal, such as substantial changes to: (1) appraised records and/or their context, (2) the creator's organizational mandate and responsibilities, (3) the creator's record-making or recordkeeping activities or systems, (4) the preserver's records preservation activities or systems and/or (5) the preserver's organizational mandate and responsibilities.

A4.3, Acquire Selected Records
To bring records selected for permanent preservation into the custody of the preserver by registering and verifying transfers, confirming the feasibility of preservation, and accessioning the records or rejecting transfers if they are inadequate.

A4.3.1, Monitor Performance of Preservation Acquisition System
To assess the efficacy of the performance of the permanent preservation acquisition system by analyzing reports on the operation of preservation activities, and issue activity

directives for acquisition activities and information on the performance of the permanent preservation selection system for use in continued maintenance of the permanent preservation system.

A4.3.2, Process Records Transfers

To register records transfers received by the designated preserver, confirm the authorization for the transfers, verify their content, confirm the authenticity of the records in the transfers and confirm the feasibility of preserving the transferred records.

A4.3.2.1, Register Transfers

To record registration information about received transfers and issue notifications of receipt to the persons transferring the records.

A4.3.2.2, Confirm Authorization for Transfers

To verify the authority for transfer of records selected for preservation and, in cases of unauthorized transfers, issue notifications of rejection of transfer to the persons transferring the records.

A4.3.2.3, Verify Content of Transfers

To determine whether transfers of records selected for preservation have been successfully transmitted (i.e., are not corrupted) and include all records and aggregates of records specified in the terms and conditions of the transfers and, in corrupted or unverified cases, issue notifications of rejection of transfer to the persons transferring the records.

A4.3.2.4, Confirm Authenticity of Records

To determine whether the assessment of the authenticity of the creator's records being transferred, which was conducted as part of the appraisal process, is still valid by verifying that the attributes relating to the records' identity and integrity have been carried forward with them along with any relevant documentation.

A4.3.2.5, Confirm Feasibility of Preservation

To confirm that the determinations of the feasibility of preservation made during the process of appraisal are still valid and, in unconfirmed cases, issue notifications of rejection of transfer to the persons transferring the records.

A4.3.3, Accession Records

To formally accept records selected for permanent preservation into custody and document transfers in accessions documentation.

A4.4, Preserve Accessioned Records

To manage information about, and the description and storage of, records acquired for permanent preservation.

A4.4.1, Manage Information About Preserved Records

To compile information about records in the permanent preservation system and about records preservation activities and to provide overall control and co-ordination of that information for use in records selection, acquisition, description, storage and access activities.

A4.4.1.1, Monitor Performance of Preservation Information System

To assess the efficacy of the performance of the permanent preservation information system by analyzing reports on the operation of preservation activities, and issue activity directives for information activities and information on the performance of the permanent preservation selection system for use in continued maintenance of the permanent preservation system.

A4.4.1.2, Compile Information for Preservation

To collect, organise and record relevant appraisal, acquisition, accession and preservation information about acquired records for their preservation, description, storage, retrieval and output.

A4.4.1.3, Update Information on Preservation Actions

To record information about actions taken to back-up, correct, update and refresh digital components of records acquired for permanent preservation or their storage.

A4.4.2, Manage Arrangement of Preserved Records

To provide overall control and co-ordination of records arrangement activities.

A4.4.3, Manage Description of Preserved Records

To provide overall control and co-ordination of records description activities, including monitoring the preservation description system, describing preserved records and developing description instruments.

A4.4.3.1, Monitor Performance of Preservation Description System

To assess the efficacy of the performance of the permanent preservation description system by analyzing reports on the operation of preservation activities, and issue activity directives for description activities and information on the performance of the permanent preservation selection system for use in continued maintenance of the permanent preservation system.

A4.4.3.2, Describe Preserved Records

To record information about the nature and make-up of records acquired for permanent preservation and about their juridical-administrative, provenancial, procedural, documentary and technological contexts, as well as information about any changes they have undergone since they were first created.

A4.4.3.3, Develop Description Instruments

To prepare tools that provide intellectual and physical control over the records in the preservation system, such as guides, inventories, indexes, repository locators and related finding aids.

A4.4.4, Manage Storage of Preserved Records

To provide overall control and co-ordination of the permanent preservation storage system and the records stored in the system by placing the records in storage, maintaining their digital components and monitoring the performance of the storage system.

A4.4.4.1, Monitor Performance of Permanent Preservation Storage System

To assess the efficacy of the performance of the permanent preservation storage system by analyzing reports on the operation of preservation activities, and issue activity directives for storage activities and information on the performance of the permanent preservation selection system for use in continued maintenance of the permanent preservation system.

A4.4.4.2, Place Preserved Records in Storage

To place the digital components of preserved records and their metadata into storage in accordance with the procedures for maintaining authentic copies of records and the actions prescribed by the preservation storage system strategies, rules and procedures and activity directives.

A4.4.4.3, Maintain Records in Permanent Preservation Storage System

To monitor the storage of preserved records and their digital components, periodically back-up the permanent preservation storage system and, as necessary, correct problems with and update the digital components, and/or refresh storage media to ensure the records in the system remain accessible, legible and intelligible over time.

A4.4.4.3.1, Monitor Preserved Records in Storage

To keep track of the condition and maintenance requirements of preserved records--more specifically, their digital components and metadata--and the media on which they are stored in the permanent preservation storage system to identify storage that needs backing-up, digital components and metadata that need correcting or updating and media that need refreshing; and to issue reports on maintenance activities.

A4.4.4.3.2, Back-up Preservation Storage System

To routinely make a copy of all digital content in the preservation storage system, including the operating system, the software applications and all digital objects in the system, for the purpose of recovery in the event of a disaster resulting in system failure or corruption, and record information about these back-up activities.

A4.4.4.3.3, Correct Problems with Preserved Records in Storage

To take the actions prescribed by the preservation storage system strategies, rules and procedures and activity directives, in accordance with the procedures for maintaining authentic copies of records, to identify and eliminate problems in storage to ensure that the records remain accessible, legible and intelligible over time; and record information about these correction activities.

A4.4.4.3.4, Update Preserved Records in Storage

To carry out conversion actions on the digital components of preserved records in storage in accordance with the procedures for maintaining authentic copies of records and the actions prescribed by the preservation storage system strategies, rules and procedures and activity directives, to ensure the records remain accessible, legible and intelligible over time (such as by migration, standardization or transformation to persistent form), and record information about these updating activities.

A4.4.4.3.5, Refresh Media for Preserved Records in Storage

To copy or transfer the digital components of preserved records in storage from one medium to another, or otherwise ensure the storage medium remains sound, in accordance with the procedures for maintaining authentic copies of records and the actions prescribed by the preservation storage system strategies, rules and procedures and activity directives, and record information about these media refreshment activities.

A4.5, Output Records

To facilitate discovery of records and/or information about records in the permanent preservation system, manage requests for preserved records and/or information about the records and monitor the performance of the permanent preservation access system.

A4.5.1, Monitor Performance of Preservation Access System

To assess the efficacy of the performance of the permanent preservation access system by analyzing reports on the operation of preservation activities, and issue activity directives for access activities and information on the performance of the permanent preservation access system for use in continued maintenance of the permanent preservation system.

A4.5.2, Facilitate Discovery of Preserved Records and/or Information

To provide authorized internal and external users with mediated access to and, as necessary, assistance in the use of, the tools and resources needed to support querying and searching for information, records and/or records aggregates in the permanent preservation system.

A4.5.3, Manage Requests for Preserved Records and/or Information

To provide overall control and co-ordination of internal and external requests for access to preserved records and/or information about the records by processing access requests, retrieving digital components for requested records and/or information, verifying retrieved components and information and providing access to retrieved records and/or information.

A4.5.3.1, Process Requests for Preserved Records and/or Information

To register access requests for preserved records and/or information, translate them, define request specifications, generate retrieval requests and account for any problems with processing access requests.

A4.5.3.1.1, Register Preservation Access Requests

To record registration information about received requests for access to preserved records and/or information about the records and issue notifications of receipt to the persons requesting the records.

A4.5.3.1.2, Retrieve Information to Process Preservation Access Requests

To gather the information, from description instruments and other preservation information, needed to process access requests for preserved records and/or information about records.

A4.5.3.1.3, Generate Preservation Retrieval Requests

To translate access requests for preserved records and/or information translated into requests to the permanent preservation storage and information systems for retrieval of the exact digital components and/or information required to fulfil the access requests.

A4.5.3.1.4, Generate Preservation Requests Specifications

To issue instructions to the preservation retrieval and access systems on how to fulfil requests for preserved records and/or information about the records based on analyses of the requests and processing information in relation to preservation retrieval and access systems' strategies, rules and procedures (including procedures for maintaining authentic copies of records) and access privileges.

A4.5.3.2, Retrieve Requested Preserved Records and/or Information

To output copies of digital components of records, information about digital components of records, rendering information about records and/or content information about records retrieved from storage in the permanent preservation system in response to retrieval requests for components and/or information and in accordance with any request specifications.

A4.5.3.3, Verify Retrieved Preserved Records and/or Information

To determine whether all components and information necessary to satisfy access requests for preserved records and/or information about the records have been received and can be processed for output and, in cases where digital components are encountered that need updating or correcting, redirect them, along with information about the problems encountered, to the maintenance function of the permanent preservation storage system for further action..

A4.5.3.4, Provide Access to Retrieved Preserved Records and/or Information

To fulfil access requests by either reconstituting the retrieved digital components of preserved records and/or information in authentic form and presenting the manifested records or information to users, or by packaging the retrieved digital components with information about how to reconstitute and present the records and/or information with the appropriate extrinsic form and issuing the packaged materials to users, and account for the success or failure of either activity.

A4.5.3.4.1, Reconstitute Preserved Records and/or Information

To link or assemble all the verified digital components of requested preserved records and/or information about preserved records as necessary to reproduce and present the records and/or information in authentic form and, if necessary, redact information to meet privacy and/or copyright requirements.

A4.5.3.4.2, Manifest Preserved Records and/or Information

To present copies of the reconstituted requested preserved records and/or requested information about the records with the appropriate extrinsic form and with information about their relationships to one another (archival bond) and, if requested, produce a Certificate of Authenticity for the records copies.

A4.5.3.4.3, Package Preserved Records and/or Information for Output

To combine the digital components of the requested preserved records and/or requested information about preserved records with information on how to reconstitute and manifest the records or information with the appropriate extrinsic form.

Chain of Preservation Model Arrow Definitions

Accession Registration Scheme
A plan for assigning a unique identifier to each accessioned records transfer.

Accessioned Records
Records that are taken into the custody of the preserver for permanent preservation.

Accounting for Preservation Access Requests
Information about successful access requests for preserved records and/or information about preserved records, including a log of the records or information provided to users, the dates when access was provided and the names of the users to whom access was provided.

Accounting for Recordkeeping Access Requests
Information about successful access requests for kept records and/or information about kept records, including a log of the records or information provided to users, the dates when access was provided and the names of the users to whom access was provided.

Analysis of Creator's Records
An analysis of the records generated by the creator in the course of its activity and of the way they were created, organized, maintained and used.

Analysis of Designated Records Preserver
An analysis of the designated records preserver's mission, organizational structure, activities, functions and existing technological, financial and human resources, and records preservation related needs and risks relevant to the identification of the framework requirements.

Analysis of Records Creator
An analysis of the record creator's mission, organizational structure, activities, functions and existing technological, financial and human resources, and records related needs and risks relevant to the identification of the framework requirements.

Appraisal Decisions
Determinations of the retention periods and disposition of records, including the terms and conditions of transfer from the creator to the preserver.

Archival Concepts, Principles and Requirements
The concepts, principles and methodologies governing the treatment of records, including the requirements for maintaining authentic copies of records.

Arranged Records
Records of a creator that have been identified as to their provenance and relationships according to the concepts and principles of archival arrangement.

Assessments of Authenticity
Documentation of the grounds for presuming the authenticity of records or, in cases of insufficient evidence to support such presumption, documentation of the verification of authenticity.

Assessments of Continuing Value
Documentation of the reasons for continuing preservation of records with regard to their capacity to serve the continuing interests of their creator and/or society.

Authentic Records
Records whose authenticity is presumed or has been verified.

Authorized Transfers
Transfers of records selected for preservation that have been submitted by persons having the authority to transfer the records.

Captured Documents
Documents made or received by the creator that are recorded and saved in the record-making system with fixed form and stable content.

Completed Records
Records, made or received by the creator, which have participated in the formal execution phase of an administrative procedure.

Completed Records Prepared for Transfer to Recordkeeping System
Executed records in the proper format for transfer to the recordkeeping system as prescribed by recordkeeping system rules and procedures and technological requirements.

Completed Records Transfers
Completed records that have been adjudged worthy of retention for future use or reference by the creator and that meet all requirements for transfer to the recordkeeping system.

Controlled Preservation Vocabulary/Thesaurus
A managed set of purposefully delimited and standardised terms, phrases and concepts used by the designated preserver to control the values of a metadata element.

Controlled Recordkeeping Vocabulary/Thesaurus
A managed set of purposefully delimited and standardised terms, phrases and concepts used by the creator to control the values of a metadata element.

Corrected Kept Records
Kept records from which problems with locating, retrieving or reconstituting their digital components and/or presenting the reconstituted records have been eliminated.

Corrected Preserved Records
Preserved records from which problems with locating, retrieving or reconstituting their digital components and/or presenting the reconstituted records have been eliminated.

Creator
An entity that generates records in the course of its activity.

Creator's Certificates of Authenticity
Attestations by the creator that one or more records are authentic.

Creator's Existing Records
Inactive, semi-active and active records of the creator, regardless of the medium and location of the records, which predate development and implementation of the chain of preservation framework and that need to be incorporated into any new record-making and recordkeeping systems.

Declared Records
Identified documents made or received by the creator that have been given a classification code based on the classification scheme and that have been registered according to the registration scheme.

Described Records
Arranged records for which information about their nature, make-up and contexts (juridical-administrative, provenancial, procedural, documentary and technological) are recorded to facilitate intellectual and physical control.

Description Instruments

Tools prepared in the course of archival description and indexing of records for the purposes of intellectual and physical control.

Design Requirements

The record-making, recordkeeping and permanent preservation needs that guide the framework design.

Disposition Rules and Procedures

The authoritative instructions governing the process of determining the transfer and destruction of kept records.

Documentation About Destroyed Records

Formal instruments documenting the destruction of kept records, including information about the quantity and characteristics of records that have been destroyed, copies of which are maintained by the creator as evidence of the activity.

Evidence for the Presumption of Authenticity

Information that has been drawn from records, from metadata related to the records and/or from their various contexts and that provides evidence to support a presumption of the authenticity of records.

Facilities

Physical space and infrastructure needed to manage the lifecycle of records.

Feasibility Reports

Assessments of whether the record elements and digital components of a given body of records proposed for preservation can be preserved given the preserver's current and anticipated preservation capabilities.

Framework Policies

Collective, high-level management principles that help guide and control development of the framework requirements.

Identified Documents

Documents made or received by the creator to which identity metadata (e.g., persons, actions and dates of compilation) have been attached.

Indexed Records

Kept records for which access points have been established using a controlled recordkeeping vocabulary to facilitate record discovery and retrieval.

Indexing Instruments

Tools that facilitate efficient and effective discovery and retrieval of kept records and/or records aggregates suited to a particular inquiry or purpose.

Information About Accessioned Records

Documentation of the provenance and custody of clearly identified sets of records for which the preserver has accepted responsibility for permanent preservation.

Information About Appraisal Decisions

Documentation explaining the justifications of appraisal decisions according to assessment of the value of records and the feasibility of their permanent preservation.

Information About Appraised Records

Documentation compiled during the appraisal process containing information about the context and content of appraised records, including information about digital components to be preserved.

Information About Available Technology
Documentation concerning the software and hardware available on the market to the creator and to the preserver.

Information About Completed Records Prepared for Transfer to Recordkeeping System
Documentation, either in the form of metadata inextricably attached to records or inextricably linked to records in record profiles, about the identity, integrity, format, form, context and other characteristics of completed records adjudged worthy of transfer to the recordkeeping system that is needed to order the records properly with respect to their relationships with each other, to maintain their authenticity and to meet recordkeeping system transfer requirements.

Information About Context
Documentation compiled about the juridical-administrative, provenancial, procedural, documentary and/or technological contexts of kept records that is not available from the records themselves, for the purpose of facilitating appraisal.

Information About Creator
Documentation concerning the records creator's mission, organizational structure, activities, and existing technological, financial and human resources, as well as information about records related needs and risks.

Information About Creator's Existing Records
Documentation about the character and extent of the records created and kept by the creator prior to developing the framework requirements.

Information About Digital Components of Kept Records in Storage
Technical documentation compiled about digital components of records in the recordkeeping storage system for the purpose of facilitating discovery of, and/or processing access requests for, records and/or information about records.

Information About Digital Components of Preserved Records in Storage
Technical documentation concerning digital components of records in the preservation storage system that is needed to facilitate discovery of, and/or process access requests for, the records and/or information about the records.

Information About Digital Components to be Preserved
Documentation about how record elements to be preserved are manifested in the electronic environment, construed for the purposes of instructing preservation activities.

Information About Executed Records
Documentation, either in the form of metadata inextricably attached to the records or inextricably linked to the records in record profiles, about the identity, integrity, format, form, context and other characteristics of executed records that is needed to order the records properly with respect to their relationships with each other and to maintain their authenticity.

Information About Implementation Problems
Documentation compiled about problems encountered during implementation of the record-making, recordkeeping, and/or permanent preservation systems for the purpose of revising the framework design process.

Information About Kept Records for Appraisal
Documentation compiled about the identity, integrity, format, form, context or other characteristics of records in the recordkeeping system for the purpose of appraising records and making appraisal decisions.

Information About Kept Records for Creation

Documentation compiled about records in the recordkeeping system for the purpose of helping direct records creation activities.

Information About Kept Records Identified for Destruction

Documentation about records in the recordkeeping system that are earmarked for destruction that is to be destroyed along with the records and/or that is used to provide information for documentation about destroyed records.

Information About Kept Records Identified for Preservation

Documentation about records in the recordkeeping system that are earmarked for transfer to the designated preserver that is needed to prepare the records in accordance with the terms and conditions of transfer.

Information About Kept Records in Storage

Documentation compiled about kept records in the recordkeeping storage system for the purpose of processing retrieval requests for records and/or information about records.

Information About Kept Records Prepared for Transfer to Preserver

Documentation about kept records and any modifications made to them in preparation for transfer to the designated preserver that is used to generate documentation about the records being transferred.

Information About Made Documents' Context

Documentation about the juridical-administrative, provenancial, procedural, documentary and/or technological context of documents made by the creator that is not available from the documents themselves and that needs to be recorded as metadata (e.g., the action or matter of the documents).

Information About Maintenance of Kept Records in Storage

Continuously updated documentation indicating the location of digital components of kept records in storage, the presence, nature and locations of recordkeeping system backups, the occurrence of storage problems, the actions taken to correct storage problems, the actions taken to update records and refresh storage media, the results of such actions and their impact, if any, on the authenticity of the records.

Information About Maintenance of Preserved Records in Storage

Continuously updated documentation indicating the location of digital components of preserved records in storage, the presence, nature and locations of permanent preservation system backups, the occurrence of storage problems, the actions taken to correct storage problems, the actions taken to update records and refresh storage media, the results of such actions--including any problems encountered--and their impact, if any, on the authenticity of the records.

Information About Outgoing Documents

Documentation about the identity, integrity, format, form, context, content or other characteristics about documents sent to external juridical or physical persons by the creator, either in the form of metadata inextricably attached to record copies of the documents retained by the creator or inextricably linked to such copies in record profiles.

Information About Preservation Storage Correction Activities

Continuously logged and updated documentation concerning actions taken to identify and eliminate problems in permanent preservation system storage, and the results of such actions, including any problems that occurred in the process and any impacts to the authenticity of preserved records and their digital components.

Information About Preservation Storage Media Refreshment Activities
Continuously logged and updated documentation concerning refreshment actions taken to ensure preservation storage media remain sound, and the results of such actions, including any problems that occurred in the process and any impacts to the authenticity of preserved records and their digital components.

Information About Preservation Storage Updating Activities
Continuously logged and updated documentation concerning conversion actions taken to ensure preserved records remain accessible, legible and intelligible over time, and the results of such actions, including any problems that occurred in the process and any impacts to the authenticity of preserved records and their digital components.

Information About Preservation System Backup Activities
Continuously logged and updated documentation concerning permanent preservation system backup and recovery activities and the results of such actions, including any problems that occurred in the process and any impacts to the authenticity of preserved records and their digital components.

Information About Preserved Records in Storage
Documentation compiled about preserved records in the permanent preservation storage system for the purpose of processing retrieval request for records and/or information about records.

Information About Preserver
Documentation concerning the designated preserver's mission, organizational structure, activities, and existing technological, financial and human resources, as well as information about records preservation-related needs and risks.

Information About Preserver's Existing Holdings
Documentation compiled about the records and aggregations of records already in the preserver's custody for the purposes of helping make valuation determinations during appraisals and helping facilitate accessioning of accruals during acquisition.

Information About Received Documents' Context
Documentation about the context of incoming documents that is not available from the documents themselves and that needs to be recorded as metadata (e.g., the action or matter of the documents).

Information About Recordkeeping Storage Correction Activities
Continuously logged and updated documentation concerning actions taken to identify and eliminate problems in recordkeeping system storage, and the results of such actions, including any impacts on the authenticity of kept records and their digital components.

Information About Recordkeeping Storage Media Refreshment Activities
Continuously logged and updated documentation concerning refreshment actions taken to ensure recordkeeping storage media remain sound, and the results of such actions, including any impacts on the authenticity of kept records and their digital components.

Information About Recordkeeping Storage Updating Activities
Continuously logged and updated documentation concerning conversion actions taken to ensure kept records remain accessible, legible and intelligible over time, and the results of such actions, including any impacts on the authenticity of kept records and their digital components.

Information About Recordkeeping System Backup Activities
Continuously logged and updated documentation concerning recordkeeping system backup and recovery activities and the results of such actions, including any impacts on the authenticity of kept records and their digital components.

Information About Reproduced Kept Records Presented to Users
Documentation about the identity, integrity, format, form, context, content or other characteristics of reproduced kept records that were presented to users to satisfy requests.

Information About Reproduced Preserved Records Presented to Users
Documentation about the identity, integrity, format, form, context, content or other characteristics of reproduced preserved records that were presented to users to satisfy requests.

Information About Reproducible Kept Records Issued to Users
Documentation about the identity, integrity, format, form, context, content or other characteristics of reproducible kept records that were issued to users to satisfy requests.

Information About Reproducible Preserved Records Issued to Users
Documentation about the identity, integrity, format, form, context, content or other characteristics of reproducible preserved records that were issued to users to satisfy requests.

Information About Retrieved Digital Components of Kept Records
Technical documentation compiled about digital components of kept records in storage for the purpose of reconstituting the requested records from the components and presenting them in authentic form to users.

Information About Retrieved Digital Components of Preserved Records
Technical documentation compiled about digital components of preserved records in storage for the purpose of reconstituting the requested records from the components and presenting them in authentic form to users.

Information About Retrieved Digital Components that Need Updating or Correcting
Documentation about retrieved digital components that cannot be reconstituted or presented in accordance with current access strategies applicable to those records.

Information About Retrieved Kept Records
Documentation compiled about retrieved kept records for the purpose of fulfilling access requests: 1) for records properly ordered with respect to their relationships with each other, or 2) for information about the identity, integrity, format, form, context, content or other characteristics of the records.

Information About Retrieved Preserved Records
Documentation compiled about retrieved preserved records for the purpose of fulfilling access requests: 1) for records properly ordered with respect to their relationships with each other, or 2) for information about the identity, integrity, format, form, context, content or other characteristics of the records.

Information About Transferred Completed Records
Documentation compiled about completed records transferred to the recordkeeping system for the purpose of: 1) establishing the identity and demonstrate the integrity of the records being transferred, 2) identifying their logical format, constituent digital components, documentary form and other recordkeeping-related characteristics, 3) properly ordering the records with respect to their relationships with each other (archival bond) and 4) placing the records in their relevant contexts (juridical-procedural, provenancial, procedural, documentary, technical).

Information About Transferred Kept Records

Documentation compiled about kept records transferred to the designated preserver for the purposes of: 1) establishing the identity and demonstrating the integrity of the records being transferred, 2) identifying their logical format, constituent digital components, documentary form and other preservation-related characteristics, 3) properly ordering the records with respect to their relationships with each other (archival bond) and 4) associating the records with their relevant contexts (juridical-procedural, provenancial, procedural, documentary, technical).

Information About Valuation Determinations

Information about the criteria used to assess the value of records and their application in a given case.

Information for Appraisal

Documentation compiled about records and their contexts for the purpose of assessing their value and authenticity.

Information for Arrangement

Documentation compiled about acquired and accessioned records and their preservation for the purpose of arranging the preserved records of a given creator.

Information for Description

Documentation compiled about acquired and accessioned records and their preservation for the purpose of describing preserved records and creating description instruments.

Information for Feasibility

Documentation compiled about records and their contexts for the purpose of determining the feasibility of their preservation.

Information for Indexing

Documentation compiled about kept records for the purpose of establishing access points and creating indexing instruments to facilitate record discovery and retrieval.

Information for Preservation

Documentation compiled about accessioned records and their elements and digital components for the purpose of facilitating preservation.

Information for Preservation Retrieval Requests

Documentation compiled about preserved records and/or information about preserved records and their digital components for the purpose of generating retrieval requests and request specifications.

Information for Recordkeeping Retrieval Requests

Documentation compiled about kept records and/or information about kept records and their digital components for the purpose of generating retrieval requests and request specifications.

Information for Storage of Kept Records

Documentation compiled about kept records and their elements and digital components for the purpose of facilitating their storage and continued maintenance.

Information for Storage of Preserved Records

Documentation compiled about preserved records and their elements and digital components for the purpose of facilitating their storage and long-term preservation.

Integrated Business and Documentary Procedures

Procedures for carrying out the creator's business that have been linked to a scheme or plan for organization of the creator's records.

Juridical system
A social group that is organized on the basis of a system of rules and that includes three components: the social group, the organizational principle of the social group and the system of binding rules recognized by the social group.

Kept Records Identified for Destruction
Records and information about records in the recordkeeping system that are identified for destruction in accordance with retention decisions.

Kept Records Identified for Preservation
Records and information about records in the recordkeeping system that are identified, in accordance with retention decisions, for transfer to the designated preserver for long-term preservation.

Kept Records in Storage
Kept records whose digital components have been placed in a storage system on digital media.

Kept Records on Media that Need Refreshing
Kept records whose digital components are stored on media that need to be refreshed to ensure the records remain accessible, legible and intelligible over time.

Kept Records on Refreshed Storage Media
Kept records and/or their digital components that have been copied or transferred to new storage media.

Kept Records Prepared for Transfer to Preserver
Records and information about records in the recordkeeping system in the proper format for transfer to the designated preserver.

Kept Records that Need Correcting
Kept records whose digital components cannot be located, retrieved, reconstituted or presented in accordance with current recordkeeping strategies applicable to those records.

Kept Records that Need Updating
Kept records whose digital components require conversion to ensure the records remain accessible, legible and intelligible over time.

Kept Records Transfers
Aggregations of kept records adjudged worthy of transfer to the designated preserver for long-term preservation and that meet all terms and conditions of transfer.

Lists of Digital Components to be Preserved
Information about the components in the electronic environment manifesting record elements that should be preserved to maintain authenticity.

Lists of Record Elements to be Preserved
Information about the extrinsic and intrinsic elements of form that need to be preserved to maintain the authenticity of records.

Made Documents
Discrete aggregations of digital information that have been compiled in a syntactic manner in accordance with the specifications of the creator's documentary forms, integrated business and documentary procedures and record-making access privileges, but which have not yet been captured (i.e., affixed to a digital medium with fixed form and stable content).

Management Policies
Formalized statements designed to provide governance and guidance in the establishment of overall framework design requirements.

Mediated Access Requests for Kept Records and/or Information

Requests from internal or external users to consult or receive copies of kept records or information about kept records in storage that have been formulated following access to and, as necessary, assistance in the use of, the tools and resources needed to support querying and searching for, and discovery of, information, records and/or records aggregates in the recordkeeping system.

Mediated Access Requests for Preserved Records and/or Information

Requests from internal or external users to consult or receive copies of preserved records or information about preserved records in storage that have been formulated following access to and, as necessary, assistance in the use of, the tools and resources needed to support querying and searching for, and discovery of, information, records and/or records aggregates in the permanent preservation system.

Need for Verification

The need to employ methods of verification of the authenticity of records as a result of there being weak evidence for the presumption of their authenticity.

Notifications of Receipt of Preservation Access Request

Formal instruments sent to the persons requesting access to preserved records and/or information about the records acknowledging that the preserver has received the request and, if needed, asking requestors to address any problems identified in registering the requests.

Notifications of Receipt of Recordkeeping Access Request

Formal instruments sent to the persons requesting access to kept records and/or information about the records acknowledging that the creator has received the request and, if needed, asking requestors to address any problems identified in registering the requests.

Notifications of Receipt of Transfer

Formal instruments sent to the creator acknowledging that the preserver has received the transfers and, if needed, asking the creator to address any problems encountered in registering the transfers.

Notifications of Rejection of Preservation Access Request

Formal instruments sent to the persons requesting access to preserved records and/or information about the records indicating that requests cannot be fulfilled because the requests are unauthorized (e.g., due to copyright restrictions), do not contain information that is sufficiently accurate, valid or complete to process the request, or are for records and/or information that cannot be located, retrieved, verified, reconstituted, manifested and/or packaged due to administrative, technical or other problems.

Notifications of Rejection of Recordkeeping Access Request

Formal instruments sent to the persons requesting access to kept records and/or information about the records indicating that requests cannot be fulfilled because the requests are unauthorized (e.g., due to access restrictions), do not contain information that is sufficiently accurate, valid or complete to process the request, or are for records and/or information that cannot be located, retrieved, verified, reconstituted, manifested and/or packaged due to administrative, technical or other problems.

Notifications of Rejection of Transfer

Formal instruments sent to the creator by the preserver indicating that transfers of records do not satisfy requirements for being accessioned or preserved, because the transfers are unauthorized, do not contain the proper records, or contain records that cannot be authenticated or whose preservation is not feasible.

Orders to Rectify Unverifiable Retrievals
Official requests issued by access activity management staff to remedy problems that resulted in retrieval of incomplete, incorrect or unprocessable digital components and/or information and, as appropriate, reattempt retrievals.

Outgoing Documents
Records that are sent to external juridical and physical persons in the course of the activities of the records creator, drafts of which are also sent to and stored in the recordkeeping system.

Permanent Preservation System
A set of rules governing the permanent intellectual and physical maintenance of records and the tools and mechanisms used to implement these rules.

Permanent Preservation System Design
The plan for the permanent preservation system outlining the selection, acquisition, description, storage, retrieval and access sub-systems.

Permanent Preservation System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the permanent preservation system.

Preservation Access System
A set of rules governing the methods and strategies for discovering, reconstituting and presenting and/or packaging retrieved records and/or information about records in the permanent preservation system and the tools and mechanisms used to implement these rules.

Preservation Access System Activity Directives
Authoritative procedural orders/instruments, issued in response to ongoing system monitoring and performance evaluations, that are intended to help direct, update and coordinate the ongoing activities of the preservation access system.

Preservation Access System Design
The plan for the access sub-system of the permanent preservation system.

Preservation Access System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the permanent preservation access sub-system in relation to the established performance requirements for the sub-system.

Preservation Access System Performance Criteria
The operational benchmarks or standards for operation of the permanent preservation access sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Preservation Access System Performance Information
Continuously logged and updated documentation concerning the ability of the permanent preservation access sub-system to fulfil is purpose and achieve its performance objectives.

Preservation Access System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the permanent preservation access sub-system to fulfil its purpose.

Preservation Access System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the permanent preservation access sub-system.

Preservation Access System Rules and Procedures
The authoritative instructions governing the operation of the permanent preservation access sub-system.

Preservation Access System Strategies
The authoritative objectives and methods governing the operation of the permanent preservation access sub-system.

Preservation Access System Technological Requirements
Specification of the hardware and software needed for the permanent preservation access sub-system.

Preservation Acquisition System
A set of rules governing the acquisition and accessioning of records transfers and the tools and mechanisms used to implement these rules.

Preservation Acquisition System Activity Directives
Authoritative procedural orders/instruments, issued in response to ongoing system monitoring and performance evaluations, that are intended to help direct, update and coordinate the ongoing activities of the preservation acquisition system.

Preservation Acquisition System Design
The plan for the acquisition sub-system of the permanent preservation system.

Preservation Acquisition System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the permanent preservation acquisition sub-system in relation to the established performance requirements for the sub-system.

Preservation Acquisition System Performance Criteria
The operational benchmarks or standards for operation of the permanent preservation acquisition sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Preservation Acquisition System Performance Information
Continuously logged and updated documentation concerning the ability of the permanent preservation acquisition sub-system to fulfil is purpose and achieve its performance objectives.

Preservation Acquisition System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the permanent preservation acquisition sub-system to fulfil its purpose.

Preservation Acquisition System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the permanent preservation acquisition sub-system.

Preservation Acquisition System Rules and Procedures
The authoritative instructions governing the operation of the permanent preservation acquisition sub-system.

Preservation Acquisition System Strategies
The authoritative objectives and methods governing the operation of the permanent preservation acquisition sub-system.

Preservation Acquisition System Technological Requirements
Specification of the hardware and software needed for the permanent preservation acquisition sub-system.

Preservation Activity Directives
Authoritative procedural orders/instruments, issued in response to ongoing system monitoring and performance evaluations, that are intended to help direct, update and coordinate the ongoing activities of the permanent preservation system.

Preservation Description System
A set of rules governing the description of preserved records and the development of description instruments and the tools and mechanisms used to implement these rules.

Preservation Description System Activity Directives
Authoritative procedural orders/instruments, issued in response to ongoing system monitoring and performance evaluations, that are intended to help direct, update and coordinate the ongoing activities of the preservation description system.

Preservation Description System Design
The plan for the description sub-system of the permanent preservation system.

Preservation Description System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the permanent preservation description sub-system in relation to the established performance requirements for the sub-system.

Preservation Description System Performance Criteria
The operational benchmarks or standards for operation of the permanent preservation description sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Preservation Description System Performance Information
Continuously logged and updated documentation concerning the ability of the permanent preservation description sub-system to fulfil is purpose and achieve its performance objectives.

Preservation Description System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the permanent preservation description sub-system to fulfil its purpose.

Preservation Description System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the permanent preservation description sub-system.

Preservation Description System Rules and Procedures
The authoritative instructions governing the operation of the permanent preservation description sub-system.

Preservation Description System Strategies
The authoritative objectives and methods governing the operation of the permanent preservation description sub-system.

Preservation Description System Technological Requirements
Specification of the hardware and software needed for the permanent preservation description sub-system.

Preservation Information System
A set of rules governing the management and maintenance of information about the operation of the permanent preservation system and about the preserved records in the system, including their digital components and the preservation actions applied to them, and the tools and mechanisms used to implement these rules.

Preservation Information System Activity Directives
Authoritative procedural orders/instruments, issued in response to ongoing system monitoring and performance evaluations, that are intended to help direct, update and coordinate the ongoing activities of the preservation information system.

Preservation Information System Design
The plan for the information sub-system of the permanent preservation system.

Preservation Information System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the permanent preservation information sub-system in relation to the established performance requirements for the sub-system.

Preservation Information System Performance Criteria
The operational benchmarks or standards for operation of the permanent preservation information sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Preservation Information System Performance Information
Continuously logged and updated documentation concerning the ability of the permanent preservation information sub-system to fulfil is purpose and achieve its performance objectives.

Preservation Information System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the permanent preservation information sub-system to fulfil its purpose.

Preservation Information System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the permanent preservation information sub-system.

Preservation Information System Rules and Procedures
The authoritative instructions governing the operation of the permanent preservation information sub-system.

Preservation Information System Strategies
The authoritative objectives and methods governing the operation of the permanent preservation Information sub-system.

Preservation Information System Technological Requirements
Specification of the hardware and software needed for the permanent preservation information sub-system.

Preservation Metadata Schemes
Lists of all necessary metadata to be recorded to ensure the identification and integrity of records preserved in the permanent preservation system.

Preservation Reporting Schemes
Plans for the systematic generation of documentation or reports of the preserver's preservation activities according to logically structured conventions, methods and procedural rules.

Preservation Retrieval System
A set of rules governing the retrieval of records, their digital components and/or information about the records and their components from the permanent preservation storage system and the tools and mechanisms used to implement these rules.

Preservation Retrieval System Design
The plan for the permanent preservation retrieval system outlining preservation retrieval rules and procedures, preservation retrieval strategies, and preservation retrieval technological requirements.

Preservation Retrieval System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the permanent preservation retrieval sub-system in relation to the established performance requirements for the sub-system.

Preservation Retrieval System Performance Criteria
The operational benchmarks or standards for operation of the permanent preservation retrieval sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Preservation Retrieval System Performance Information
Continuously logged and updated documentation concerning the ability of the permanent preservation retrieval sub-system to fulfil is purpose and achieve its performance objectives.

Preservation Retrieval System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the permanent preservation retrieval sub-system to fulfil its purpose.

Preservation Retrieval System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the permanent preservation retrieval sub-system.

Preservation Retrieval System Rules and Procedures
The authoritative instructions governing the operation of the permanent preservation retrieval sub-system.

Preservation Retrieval System Strategies
The authoritative objectives and methods governing the operation of the permanent preservation retrieval sub-system.

Preservation Retrieval System Technological Requirements
Specification of the hardware and software needed for the permanent preservation retrieval sub-system.

Preservation Selection System
A set of rules governing the appraisal of kept records and the tools and mechanisms used to implement these rules.

Preservation Selection System Activity Directives
Authoritative procedural orders/instruments, issued in response to ongoing system monitoring and performance evaluations, that are intended to help direct, update and coordinate the ongoing activities of the preservation selection system.

Preservation Selection System Design
The plan for the selection sub-system of the permanent preservation system.

Preservation Selection System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the permanent preservation selection sub-system in relation to the established performance requirements for the sub-system.

Preservation Selection System Performance Criteria
The operational benchmarks or standards for operation of the permanent preservation selection sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Preservation Selection System Performance Information
Continuously logged and updated documentation concerning the ability of the preservation selection sub-system to fulfil its purpose and achieve its performance objectives.

Preservation Selection System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the permanent preservation selection sub-system to fulfil its purpose.

Preservation Selection System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the permanent preservation selection sub-system.

Preservation Selection System Rules and Procedures
The authoritative instructions governing the operation of the permanent preservation selection sub-system.

Preservation Selection System Strategies
The authoritative objectives and methods governing the operation of the permanent preservation selection sub-system.

Preservation Selection System Technological Requirements
Specification of the hardware and software needed for the permanent preservation selection sub-system.

Preservation Storage System
A set of rules governing the storage of records, their digital components and/or information about the records and components in the permanent preservation system and the tools and mechanisms used to implement these rules.

Preservation Storage System Activity Directives
Authoritative procedural orders/instruments, issued in response to ongoing system monitoring and performance evaluations, that are intended to help direct, update and coordinate the ongoing activities of the preservation storage system.

Preservation Storage System Design
The plan for the storage sub-system of the permanent preservation system.

Preservation Storage System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the permanent preservation storage sub-system in relation to the established performance requirements for the sub-system.

Preservation Storage System Performance Criteria
The operational benchmarks or standards for operation of the permanent preservation storage sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Preservation Storage System Performance Information
Continuously logged and updated documentation concerning the ability of the permanent preservation storage sub-system to fulfil is purpose and achieve its performance objectives.

Preservation Storage System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the permanent preservation storage sub-system to fulfil its purpose.

Preservation Storage System Policies

The collective, high-level management principles that guide and control development, implementation and execution of the permanent preservation storage sub-system.

Preservation Storage System Rules and Procedures

The authoritative instructions governing the operation of the permanent preservation storage sub-system.

Preservation Storage System Strategies

The authoritative objectives and methods governing the operation of the permanent preservation storage sub-system.

Preservation Storage System Technological Requirements

Specification of the hardware and software needed for the permanent preservation storage sub-system.

Preservation Storage that Needs Backing Up

All software applications and digital objects in the preservation storage system that need backing up as specified by permanent preservation storage system strategies.

Preservation System Access Privileges

The authority to compile, annotate, read, retrieve, transfer, and/or destroy records in the preservation system, granted to officers and employees of the entity responsible for preservation.

Preservation System Administrative Infrastructure

A comprehensive, integrated set of administrative policies, strategies, rules and procedures, and instruments that support preservation activities and enable the permanent preservation system to meet its functional requirements.

Preservation System Backup

A copy of all digital content in the preservation storage system.

Preservation System Functional Infrastructure Design

The comprehensive, integrated design for the permanent preservation system and each of its records information, selection, acquisition, description, storage, retrieval and access sub-systems.

Preservation System Functional Requirements

The comprehensive and integrated performance, monitoring and technological requirements for the permanent preservation system.

Preservation System Instruments

The administrative tools that support the preservation of records in the permanent preservation system, such as preservation metadata schemes, records transfer and accession registration schemes and controlled preservation vocabularies and thesauri.

Preservation System Monitoring Requirements

The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the permanent preservation system in relation to the established performance requirements for the system.

Preservation System Performance Criteria

The operational benchmarks or standards for operation of the permanent preservation system against which the continuing performance and adequacy of all activities, functions, processes, sub-systems and structures within the system are measured.

Preservation System Performance Information
Information about the ability of the individual components of the permanent preservation system to fulfil their purposes.

Preservation System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the permanent preservation system to fulfil its purpose.

Preservation System Rules and Procedures
The authoritative instructions governing the operation of the permanent preservation system.

Preservation System Strategies
The authoritative objectives and methods governing the operation of the permanent preservation system.

Preservation System Technological Requirements
Specification of the hardware and software needed for the permanent preservation system.

Preserved Records in Storage
Preserved records whose digital components have been placed in a storage system on digital media.

Preserved Records on Media that Need Refreshing
Preserved records whose digital components are stored on media that need to be refreshed to ensure the records remain accessible, legible and intelligible over time.

Preserved Records on Refreshed Storage Media
Preserved records and/or their digital components that have been copied or transferred to new storage media.

Preserved Records that Need Correcting
Preserved records whose digital components cannot be located, retrieved, reconstituted or presented in accordance with current preservation strategies applicable to those records.

Preserved Records that Need Updating
Preserved records whose digital components require conversion to ensure the records remain accessible, legible and intelligible over time.

Preserver
The entity responsible for managing the permanent preservation of records.

Preserver's Certificates of Authenticity
Attestations by the preserver that one or more records are authentic.

Preserver's Mission
For an archival institution or program: the jurisdiction, mandate, functions, and requirements to preserve the appraised records; for an individual: the goals, purposes, objectives, and related business needs to preserve selected records.

Procedures for Assessing Authenticity of Records
Authoritative procedural orders designed to facilitate evaluation of the authenticity of the creator's records during appraisal and/or acquisition of the records by the designated preserver.

Procedures for Ensuring the Accuracy of Records
Authoritative procedural orders designed to ensure that records are created accurate in the record-making system.

Procedures for Ensuring the Reliability of Records
Authoritative procedural orders designed to ensure that records are created reliable in the record-making system.

Procedures for Maintaining Authentic Copies of Records
Authoritative procedural orders outlining pre-established requirements for maintaining authentic copies of the creator's records in the custody of the designated preserver.

Procedures for Maintaining Authentic Records
Authoritative procedural orders designed to ensure that records maintain their identity and integrity as they are managed and maintained in the recordkeeping system.

Received Documents
Documents received by the creator from external juridical or physical persons.

Recommendations of Need to Update Appraisal Decisions
Instructions to revise appraisal decisions as a result of substantial changes in appraised records and their context, or as a result of substantial changes to a creator's organizational mandate and responsibilities and/or its record-making or recordkeeping activities or systems.

Recommended Framework Revisions
Suggestions on revising the framework design based on assessments of performance information of the record-making, recordkeeping, and permanent preservation systems.

Reconstituted Kept Records and/or Information
The linked or reassembled digital components of, and/or information about, kept records retrieved from storage for the purpose of reproducing and presenting the requested records and/or information to users.

Reconstituted Preserved Records and/or Information
The linked or reassembled digital components of, and/or information about, preserved records retrieved from storage for the purpose of reproducing and presenting the requested records and/or information to users.

Record Copies of Outgoing Documents
Drafts or record copies of documents sent to external juridical or physical persons, which are also sent to the recordkeeping system.

Record Copies of Redacted Kept Records Issued to Users
Record copies of kept records and/or information issued to users that were redacted to meet privacy and/or copyright requirements.

Record Copies of Redacted Preserved Records Issued to Users
Record copies of preserved records and/or information issued to users that were redacted to meet privacy and/or copyright requirements.

Record Descriptions
Descriptive information about preserved records, including their nature, make-up and contexts (juridical-administrative, provenancial, procedural, documentary and technological) that is recorded to facilitate intellectual and physical control of the records and, together with description instruments, to facilitate discovery.

Record Profile Schemes
Plans for the systematic generation of digital forms designed to contain the attributes of records that attest to their identity and integrity, and which are generated when users create, send and/or close records, are updated when users subsequently modify or annotate completed records, and remain inextricably linked to the records for the entire period of their existence while in the custody of the creator.

Recordkeeping Access System

A set of rules governing the methods and strategies for discovering, reconstituting and presenting and/or packaging retrieved records and/or information about records in the recordkeeping system and the tools and mechanisms used to implement these rules.

Recordkeeping Access System Activity Directives

Authoritative procedural orders/instruments, issued in response to ongoing system monitoring and performance evaluations, that are intended to help direct, update and coordinate the ongoing activities of the recordkeeping access system.

Recordkeeping Access System Design

The plan for the access sub-system of the recordkeeping system.

Recordkeeping Access System Monitoring Requirements

The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the recordkeeping access sub-system in relation to the established performance requirements for the sub-system.

Recordkeeping Access System Performance Criteria

The operational benchmarks or standards for operation of the recordkeeping access sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Recordkeeping Access System Performance Information

Continuously logged and updated documentation concerning the ability of the recordkeeping access sub-system to fulfil is purpose and achieve its performance objectives.

Recordkeeping Access System Performance Requirements

The operational and administrative specifications for measuring the continuing ability of the recordkeeping access sub-system to fulfil its purpose.

Recordkeeping Access System Policies

The collective, high-level management principles that guide and control development, implementation and execution of the recordkeeping access sub-system.

Recordkeeping Access System Rules and Procedures

The authoritative instructions governing the operation of the recordkeeping access sub-system.

Recordkeeping Access System Strategies

The authoritative objectives and methods governing the operation of the recordkeeping access sub-system.

Recordkeeping Access System Technological Requirements

Specification of the hardware and software needed for the recordkeeping access sub-system.

Recordkeeping Activity Directives

Authoritative procedural orders/instruments, issued in response to ongoing system monitoring and performance evaluations, that are intended to help direct, update and coordinate the ongoing activities of the recordkeeping system.

Recordkeeping Classification Scheme

A plan for the systematic identification and arrangement of the creator's business activities and records into categories according to logically structured conventions, methods and procedural rules.

Recordkeeping Disposition System
A set of rules governing the disposition of kept records and the tools and mechanisms used to implement these rules.

Recordkeeping Disposition System Activity Directives
Authoritative procedural orders/instruments, issued in response to ongoing system monitoring and performance evaluations, that are intended to help direct, update and coordinate the ongoing activities of the recordkeeping disposition system.

Recordkeeping Disposition System Design
The plan for the disposition sub-system of the recordkeeping system.

Recordkeeping Disposition System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the recordkeeping disposition sub-system in relation to the established performance requirements for the sub-system.

Recordkeeping Disposition System Performance Criteria
The operational benchmarks or standards for operation of the recordkeeping disposition sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Recordkeeping Disposition System Performance Information
Continuously logged and updated documentation concerning the ability of the recordkeeping disposition sub-system to fulfil is purpose and achieve its performance objectives.

Recordkeeping Disposition System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the recordkeeping disposition sub-system to fulfil its purpose.

Recordkeeping Disposition System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the recordkeeping disposition sub-system.

Recordkeeping Disposition System Rules and Procedures
The authoritative instructions governing the operation of the recordkeeping disposition sub-system.

Recordkeeping Disposition System Strategies
The authoritative objectives and methods governing the operation of the recordkeeping disposition sub-system.

Recordkeeping Disposition System Technological Requirements
Specification of the hardware and software needed for the recordkeeping disposition sub-system.

Recordkeeping Indexing System
A set of rules governing the indexing of kept records and the tools and mechanisms used to implement these rules.

Recordkeeping Indexing System Activity Directives
Authoritative procedural orders/instruments, issued in response to ongoing system monitoring and performance evaluations, that are intended to help direct, update and coordinate the ongoing activities of the recordkeeping indexing sub-system.

Recordkeeping Indexing System Design
The plan for the indexing sub-system of the recordkeeping system.

Recordkeeping Indexing System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the recordkeeping indexing sub-system in relation to the established performance requirements for the sub-system.

Recordkeeping Indexing System Performance Criteria
The operational benchmarks or standards for operation of the recordkeeping indexing sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Recordkeeping Indexing System Performance Information
Continuously logged and updated documentation concerning the ability of the recordkeeping indexing sub-system to fulfil is purpose and achieve its performance objectives.

Recordkeeping Indexing System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the recordkeeping indexing sub-system to fulfil its purpose.

Recordkeeping Indexing System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the recordkeeping indexing sub-system.

Recordkeeping Indexing System Rules and Procedures
The authoritative instructions governing the operation of the recordkeeping indexing sub-system.

Recordkeeping Indexing System Strategies
The authoritative objectives and methods governing the operation of the recordkeeping indexing sub-system.

Recordkeeping Indexing System Technological Requirements
Specification of the hardware and software needed for the recordkeeping indexing sub-system.

Recordkeeping Information System
A set of rules governing the management and maintenance of information about the operation of the recordkeeping system and about the kept records in the system, including their digital components and metadata and the recordkeeping actions applied to them, and the tools and mechanisms used to implement these rules.

Recordkeeping Information System Design
The plan for the information sub-system of the recordkeeping system.

Recordkeeping Information System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the recordkeeping information sub-system in relation to the established performance requirements for the sub-system.

Recordkeeping Information System Performance Criteria
The operational benchmarks or standards for operation of the recordkeeping information sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Recordkeeping Information System Performance Information
Continuously logged and updated documentation concerning the ability of the recordkeeping information sub-system to fulfil is purpose and achieve its performance objectives.

Recordkeeping Information System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the recordkeeping information sub-system to fulfil its purpose.

Recordkeeping Information System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the recordkeeping information sub-system.

Recordkeeping Information System Rules and Procedures
The authoritative instructions governing the operation of the recordkeeping information sub-system.

Recordkeeping Information System Strategies
The authoritative objectives and methods governing the operation of the recordkeeping information sub-system.

Recordkeeping Information System Technological Requirements
Specification of the hardware and software needed for the recordkeeping information sub-system.

Recordkeeping Metadata Schemes
Lists of all necessary metadata to be recorded to ensure the identification and integrity of records maintained in the recordkeeping system.

Recordkeeping Registration Scheme
A plan for assigning a unique identifier to each record in the recordkeeping system.

Recordkeeping Reporting Schemes
Plans for the systematic generation of documentation or reports of the creator's recordkeeping activities according to logically structured conventions, methods and procedural rules.

Recordkeeping Retrieval System
A set of rules governing searching and finding records and/or information about records in a recordkeeping system, and the tools and mechanisms used to implement these rules.

Recordkeeping Retrieval System Design
The plan for the retrieval sub-system of the recordkeeping system.

Recordkeeping Retrieval System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the recordkeeping retrieval sub-system in relation to the established performance requirements for the sub-system.

Recordkeeping Retrieval System Performance Criteria
The operational benchmarks or standards for operation of the recordkeeping retrieval sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Recordkeeping Retrieval System Performance Information
Continuously logged and updated documentation concerning the ability of the recordkeeping retrieval sub-system to fulfil is purpose and achieve its performance objectives.

Recordkeeping Retrieval System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the recordkeeping retrieval sub-system to fulfil its purpose.

Recordkeeping Retrieval System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the recordkeeping retrieval sub-system.

Recordkeeping Retrieval System Rules and Procedures
The authoritative instructions governing the operation of the recordkeeping retrieval sub-system.

Recordkeeping Retrieval System Strategies
The authoritative objectives and methods governing the operation of the recordkeeping retrieval sub-system.

Recordkeeping Retrieval System Technological Requirements
Specification of the hardware and software needed for the recordkeeping retrieval sub-system.

Recordkeeping Storage System
A set of rules governing the storage of records, their digital components and/or information about the records and components in the recordkeeping system and the tools and mechanisms used to implement these rules.

Recordkeeping Storage System Activity Directives
Authoritative procedural orders/instruments, issued in response to ongoing system monitoring and performance evaluations, that are intended to help direct, update and coordinate the ongoing activities of the recordkeeping storage system.

Recordkeeping Storage System Design
The plan for the recordkeeping storage system outlining a set of rules governing the storage of records and/or information about records in a recordkeeping system, recordkeeping storage strategies, and recordkeeping storage technological requirements.

Recordkeeping Storage System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the recordkeeping storage sub-system in relation to the established performance requirements for the sub-system.

Recordkeeping Storage System Performance Criteria
The operational benchmarks or standards for operation of the recordkeeping storage sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Recordkeeping Storage System Performance Information
Continuously logged and updated documentation concerning the ability of the recordkeeping storage sub-system to fulfil is purpose and achieve its performance objectives.

Recordkeeping Storage System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the recordkeeping storage sub-system to fulfil its purpose.

Recordkeeping Storage System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the recordkeeping storage sub-system.

Recordkeeping Storage System Rules and Procedures
The authoritative instructions governing the operation of the recordkeeping storage sub-system.

Recordkeeping Storage System Strategies
The authoritative objectives and methods governing the operation of the recordkeeping storage sub-system.

Recordkeeping Storage System Technological Requirements
Specification of the hardware and software needed for the recordkeeping storage sub-system.

Recordkeeping Storage that Needs Backing Up
All software applications and digital objects in the recordkeeping storage system that need backing up as specified by recordkeeping storage system strategies.

Recordkeeping System
A set of rules governing the storage, use, maintenance and disposition of records and/or information about records and the tools and mechanisms used to implement these rules..

Recordkeeping System Access Privileges
The authority to annotate, read, retrieve, transfer and/or destroy records in the recordkeeping system, granted to officers and employees of the creator.

Recordkeeping System Administrative Infrastructure
A comprehensive, integrated set of administrative policies, strategies, rules and procedures, and instruments that support recordkeeping activities and enable the recordkeeping system to meet its functional requirements.

Recordkeeping System Backup
A copy of all digital content in the recordkeeping storage system.

Recordkeeping System Design
The plan for the recordkeeping system outlining the recordkeeping metadata schemes, classification scheme, retention schedule, registration scheme, recordkeeping retrieval system, recordkeeping technological requirements, recordkeeping access privileges, and procedures for maintaining authentic records.

Recordkeeping System Functional Infrastructure Design
The comprehensive, integrated design for the recordkeeping system and each of its records information, storage, retrieval, access and disposition sub-systems.

Recordkeeping System Functional Requirements
The comprehensive and integrated performance, monitoring and technological requirements for the recordkeeping system.

Recordkeeping System Instruments
The administrative tools that support the maintenance of records in the recordkeeping system, such as recordkeeping metadata schemes, records registration and classification schemes, a retention schedule and controlled recordkeeping vocabularies and thesauri.

Recordkeeping System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the recordkeeping system in relation to the established performance requirements for the system.

Recordkeeping System Performance Criteria
The operational benchmarks or standards for operation of the recordkeeping system against which the continuing performance and adequacy of all activities, functions, processes and structures within the system are measured.

Recordkeeping System Performance Information
Continuously logged and updated documentation about the ability of the individual components of the recordkeeping system to fulfil their purposes.

Recordkeeping System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the recordkeeping system to fulfil its purpose.

Recordkeeping System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the recordkeeping system.

Recordkeeping System Rules and Procedures
The authoritative instructions governing the operation of the recordkeeping system.

Recordkeeping System Strategies
The authoritative objectives and methods governing the operation of the recordkeeping system.

Recordkeeping System Technological Requirements
Specification of the hardware and software needed for the recordkeeping system.

Record-making Access Privileges
The authority to compile, annotate, read, retrieve, transfer and/or destroy records in the record-making system, granted to officers and employees of the creator.

Record-making Activity Directives
Authoritative procedural orders/instruments intended to facilitate effective, co-ordinated and responsive record-making activities.

Record-making Capture System Design
The plan for the capture sub-system of the record-making system.

Record-making Capture System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the record-making capture sub-system in relation to the established record-making requirements for the sub-system.

Record-making Capture System Performance Criteria
The operational benchmarks or standards for operation of the record-making capture sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Record-making Capture System Performance Information
Continuously logged and updated documentation concerning the ability of the record-making capture sub-system to fulfil is purpose and achieve its performance objectives.

Record-making Capture System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the record-making capture sub-system to fulfil its purpose.

Record-making Capture System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the record-making capture sub-system.

Record-making Capture System Rules and Procedures
The authoritative instructions governing the operation of the record-making capture sub-system.

Record-making Capture System Strategies
The authoritative objectives and methods governing the operation of the record-making capture sub-system.

Record-making Capture System Technological Requirements
Specification of the hardware and software needed for the record-making capture sub-system.

Record-making Declaration System Design
The plan for the declaration sub-system of the record-making system.

Record-making Declaration System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the record-making declaration sub-system in relation to the established record-making requirements for the sub-system.

Record-making Declaration System Performance Criteria
The operational and administrative specifications for measuring the continuing ability of the record-making declaration sub-system to fulfil its purpose.

Record-making Declaration System Performance Information
Continuously logged and updated documentation concerning the ability of the record-making declaration sub-system to fulfil is purpose and achieve its performance objectives.

Record-making Declaration System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the record-making declaration sub-system to fulfil its purpose.

Record-making Declaration System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the record-making declaration sub-system.

Record-making Declaration System Rules and Procedures
The authoritative instructions governing the operation of the record-making declaration sub-system.

Record-making Declaration System Strategies
The authoritative objectives and methods governing the operation of the record-making declaration sub-system.

Record-making Declaration System Technological Requirements
Specification of the hardware and software needed for the record-making declaration sub-system.

Record-making Execution System Design
The plan for the execution sub-system of the record-making system.

Record-making Execution System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the record-making execution sub-system in relation to the established record-making requirements for the sub-system.

Record-making Execution System Performance Criteria
The operational benchmarks or standards for operation of the record-making execution sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Record-making Execution System Performance Information
Continuously logged and updated documentation concerning the ability of the record-making execution sub-system to fulfil is purpose and achieve its performance objectives.

Record-making Execution System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the record-making execution sub-system to fulfil its purpose.

Record-making Execution System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the record-making execution sub-system.

Record-making Execution System Rules and Procedures
The authoritative instructions governing the operation of the record-making execution sub-system.

Record-making Execution System Strategies
The authoritative objectives and methods governing the operation of the record-making execution sub-system.

Record-making Execution System Technological Requirements
Specification of the hardware and software needed for the record-making execution sub-system.

Record-making Identification System Design
The plan for the identification sub-system of the record-making system.

Record-making Identification System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the record-making identification sub-system in relation to the established record-making requirements for the sub-system.

Record-making Identification System Performance Criteria
The operational and administrative specifications for measuring the continuing ability of the record-making identification sub-system to fulfil its purpose.

Record-making Identification System Performance Information
Continuously logged and updated documentation concerning the ability of the record-making identification sub-system to fulfil is purpose and achieve its performance objectives.

Record-making Identification System Performance Requirements
The operational benchmarks or standards for operation of the record-making identification sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Record-making Identification System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the record-making identification sub-system.

Record-making Identification System Rules and Procedures
The authoritative instructions governing the operation of the record-making identification sub-system.

Record-making Identification System Strategies
The authoritative objectives and methods governing the operation of the record-making identification sub-system.

Record-making Identification System Technological Requirements
Specification of the hardware and software needed for the record-making identification sub-system.

Record-making Metadata Schemes
Lists of all necessary record-making metadata to be recorded to ensure the reliability, accuracy, identification and integrity of records created in the record-making system.

Record-making Reporting Schemes
Plans for the systematic generation of documentation or reports of the creator's record-making activities according to logically structured conventions, methods and procedural rules.

Record-making System
A set of rules governing the making of records, and the tools and mechanisms used to implement these rules.

Record-making System Administrative Infrastructure
A comprehensive, integrated set of administrative policies, strategies, rules and procedures, and instruments that support record-making activities and enable the record-making system to meet its functional requirements.

Record-making System Design
The plan for the record-making system outlining the integrated business and documentary procedures, record-making metadata schemes, records forms, record-making technological requirements, and record-making access privileges.

Record-making System Functional Infrastructure Design
The comprehensive, integrated design for the record-making system and each of its documents and records capture, identification, declaration, execution and transfer sub-systems.

Record-making System Functional Requirements
The comprehensive and integrated performance, monitoring and technological requirements for the record-making system.

Record-making System Instruments
The administrative tools that support the preservation of records in the record-making system, such as record-making metadata schemes and records forms.

Record-making System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the record-making system in relation to the established record-making system requirements.

Record-making System Performance Criteria
The operational benchmarks or standards for operation of the record-making system against which the continuing performance and adequacy of all activities, functions, processes and structures within the system are measured.

Record-making System Performance Information
Continuously logged and updated documentation concerning the ability of the individual components of the record-making system to fulfil their purposes.

Record-making System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the record-making system to fulfil its purpose.

Record-making System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the record-making system.

Record-making System Rules and Procedures
The authoritative instructions governing the operation of the record-making system.

Record-making System Strategies
The authoritative objectives and methods governing the operation of the record-making system.

Record-making System Technological Requirements
Specification of the hardware and software needed for the record-making system.

Record-making Transfer System Activity Directives
Authoritative procedural orders/instruments intended to facilitate effective, co-ordinated and responsive record-making transfer system activities.

Record-making Transfer System Design
The plan for the transfer sub-system of the record-making system.

Record-making Transfer System Monitoring Requirements
The operational and administrative conditions that need to be established to facilitate ongoing assessment of the operation of the record-making transfer sub-system in relation to the established record-making requirements for the sub-system.

Record-making Transfer System Performance Criteria
The operational benchmarks or standards for operation of the record-making transfer sub-system against which the continuing performance and adequacy of all activities, functions, processes and structures within the sub-system are measured.

Record-making Transfer System Performance Information
Continuously logged and updated documentation concerning the ability of the record-making transfer sub-system to fulfil is purpose and achieve its performance objectives.

Record-making Transfer System Performance Requirements
The operational and administrative specifications for measuring the continuing ability of the record-making transfer sub-system to fulfil its purpose.

Record-making Transfer System Policies
The collective, high-level management principles that guide and control development, implementation and execution of the record-making transfer sub-system.

Record-making Transfer System Rules and Procedures
The authoritative instructions governing the operation of the record-making transfer sub-system.

Record-making Transfer System Strategies
The authoritative objectives and methods governing the operation of the record-making transfer sub-system.

Record-making Transfer System Technological Requirements
Specification of the hardware and software needed for the record-making transfer sub-system.

Records Forms
Specifications of the documentary forms for the various types of records of the creator.

Records Manager
Person responsible for the management of active and semi-active records of a creator. The role of a records manager should be that of a trusted records officer.

Records to be Accessioned
Records in transfers that have been registered, are authorized, meet the terms and conditions of transfer, and meet the preserver's feasibility requirements, and consequently can be accessioned for permanent preservation.

Records Transfer Registration Scheme
A plan for assigning a unique identifier to each received records transfer.

Registered Preservation Access Requests
Access requests for preserved records and/or information about the records that have been received from internal and external users and registered by the preserver.

Registered Recordkeeping Access Requests
Access requests for kept records and/or information about the records that have been received from internal and external users and registered by the records manager.

Registered Transfers
Transfers of records selected for preservation that have been received from the creator and registered by the preserver.

Reports on Operation of Permanent Preservation Activity
Documentation concerning the efficacy of the operation of activities of the permanent preservation system.

Reports on Operation of Recordkeeping Activity
Documentation concerning the efficacy of the operation of activities of the recordkeeping system.

Reports on Operation of Record-making Activity
Documentation concerning the efficacy of the operation of the activities of the record-making system.

Reproduced Kept Records Presented to Users
Authentic representations or other versions of kept records reconstituted from their digital components.

Reproduced Preserved Records Presented to Users
Authentic representations or other versions of records reconstituted from their digital components.

Reproducible Kept Records Issued to Users
The digital components of kept records together with the technical information or software necessary to reproduce them from the digital components.

Reproducible Preserved Records Issued to Users
The digital components of preserved records together with the technical information necessary to reproduce them from the digital components.

Request Specifications for Kept Records
Instructions to the recordkeeping retrieval and access systems on how to fulfil requests for digital components of kept records and/or information about kept records in storage.

Request Specifications for Preserved Records
Instructions to the preservation retrieval and access systems on how to fulfil requests for digital components of preserved records and/or information about preserved records in storage.

Requests for Updated Information About Creator
Requests for updated information concerning any significant changes to the designated preserver's mission, organizational structure, activities, and existing technological, financial and human resources, as well as information about any significant changes to the preserver's preservation-related needs and risks.

Retention Schedule
A document providing description of records series and/or classes and specifying their authorized dispositions.

Retrieval Requests for Kept Records and/or Information
Access requests for kept records and/or information translated into requests to the recordkeeping storage and information systems for retrieval of the exact digital components and/or information required to fulfil the access requests.

Retrieval Requests for Preserved Records and/or Information

Access requests for preserved records and/or information translated into requests to the permanent preservation storage and information systems for retrieval of the exact digital components and/or information required to fulfil the access requests.

Retrieved Digital Components of Kept Records

The digital components of kept records retrieved from storage in response to requests.

Retrieved Digital Components of Kept Records that Need Updating or Correcting

Digital components of kept records that cannot be reconstituted or presented in accordance with current recordkeeping access strategies applicable to those records.

Retrieved Digital Components of Preserved Records

The digital components of preserved records retrieved from storage in response to requests.

Retrieved Digital Components of Preserved Records that Need Updating or Correcting

Digital components of preserved records that cannot be reconstituted or presented in accordance with current preservation access strategies applicable to those records.

Retrieved Information About Kept Records

Documentation about the identity, integrity, format, form, context, content or other characteristics of kept records retrieved from the recordkeeping information sub-system and/or, as necessary, through examination of the records themselves, in response to requests.

Retrieved Information About Preserved Records

Documentation about the identity, integrity, format, form, context, content or other characteristics of preserved records retrieved from the permanent preservation information sub-system and/or, as necessary, through examination of the records themselves, in response to requests.

Standards

Sets of rules or guidelines co-operatively adhered to by peer entities.

State of Technology

The availability and/or capability of technology at any given time.

Technical Description of Records Proposed for Preservation

Information about the technical components and requirements of records proposed for preservation that is necessary for helping to determine the feasibility of preserving the records.

Terms and Conditions of Transfer

Formal instruments that identify in archival and technological terms digital records to be transferred, together with relevant documentation, and that identifies the medium and format of transfers, when the transfers will occur, and the parties to the transfers.

Tools

Information, technology and other equipment and supplies used to manage the lifecycle of records.

Transfer Documentation for Preservation

Formal instruments indicating the entity transferring records, the contents of transfers and the terms and conditions governing the transfers, copies of which are sent to the preserver and maintained by the creator as evidence of the transaction.

Transfer Documentation for Recordkeeping

Formal instruments indicating the entity transferring records, the contents of transfers and, as necessary, information about the records being transferred for the purposes of maintaining the records in the recordkeeping system and for providing evidence of the transaction.

Unmediated Access Requests for Kept Records and/or Information

Requests from external users to consult or receive copies of kept records and/or information about kept records that were formulated without the assistance of records management staff or access to record indexing instruments.

Unmediated Access Requests for Preserved Records and/or Information

Requests from external users to consult or receive copies of preserved records and/or information about preserved records that were formulated without the assistance of archives staff or access to record descriptions or other formal description instruments.

Updated Information About Creator

Updated information concerning significant changes to the records creator's mission, organizational structure, activities, and existing technological, financial and human resources, as well as to the creator's records-related needs and risks.

Updated Information for Preservation

Updated information about records and their elements and components that is needed for preservation.

Updated Kept Records

Kept records whose digital components have been converted or updated.

Updated Preserved Records

Preserved records whose digital components have been converted or updated.

Valuation Determinations

Decisions concerning the overall value of appraised records in relation to assessments of their authenticity and capacity to serve the continuing interests of their creator and/or society, as well as their suitability and relevance in relation to the preserver's mission and existing holdings.

Verified Information About Retrieved Digital Components of Kept Records

Technical documentation compiled about digital components of kept records retrieved from storage that has been verified to ensure that the information received is correct (i.e., pertains to the requested records), complete and sufficient to enable the records to be reconstituted in authentic form from the retrieved components.

Verified Information About Retrieved Digital Components of Preserved Records

Technical documentation compiled about digital components of preserved records retrieved from storage that has been verified to ensure that the information received is correct (i.e., pertains to the requested records), complete and sufficient to enable the records to be reconstituted in authentic form from the retrieved components.

Verified Information About Retrieved Kept Records

Documentation compiled about kept records retrieved from storage that has been verified to ensure that the information received is correct (i.e., pertains to the requested records), complete and sufficient to fulfil access requests: 1) for records properly ordered with respect to their relationships with each other, or 2) for information about the identity, integrity, format, form, context, content or other characteristics of the records.

Verified Information About Retrieved Preserved Records

Documentation compiled about preserved records retrieved from storage that has been verified to ensure that the information received is correct (i.e., pertains to the requested records), complete and sufficient to fulfil access requests: 1) for records properly ordered with respect to their relationships with each other, or 2) for information about the identity, integrity, format, form, context, content or other characteristics of the records.

Verified Retrieved Digital Components of Kept Records

The aggregations of digital components of kept records retrieved from storage in response to requests, which have been verified to ensure that all requested components have been received.

Verified Retrieved Digital Components of Preserved Records

The aggregations of digital components of preserved records retrieved from storage in response to requests, which have been verified to ensure that all requested components have been received.

Verified Retrieved Information About Kept Records

Documentation about the identity, integrity, format, form, context, content or other characteristics of kept records retrieved from the recordkeeping information sub-system and/or, as necessary, through examination of the records themselves, which has been verified to ensure that the information received is correct (i.e., pertains to the requested records), complete and sufficient to satisfy access requests for information about kept records.

Verified Retrieved Information About Preserved Records

Documentation about the identity, integrity, format, form, context, content or other characteristics of preserved records retrieved from the permanent preservation information sub-system and/or, as necessary, through examination of the records themselves, which has been verified to ensure that the information received is correct (i.e., pertains to the requested records), complete and sufficient to satisfy access requests for information about preserved records.

Verified Transfers

Transfers of records selected for preservation that have been successfully received from the creator (i.e., were not corrupted in transmission) and include all records and aggregates of records specified in the terms and conditions of the transfers.

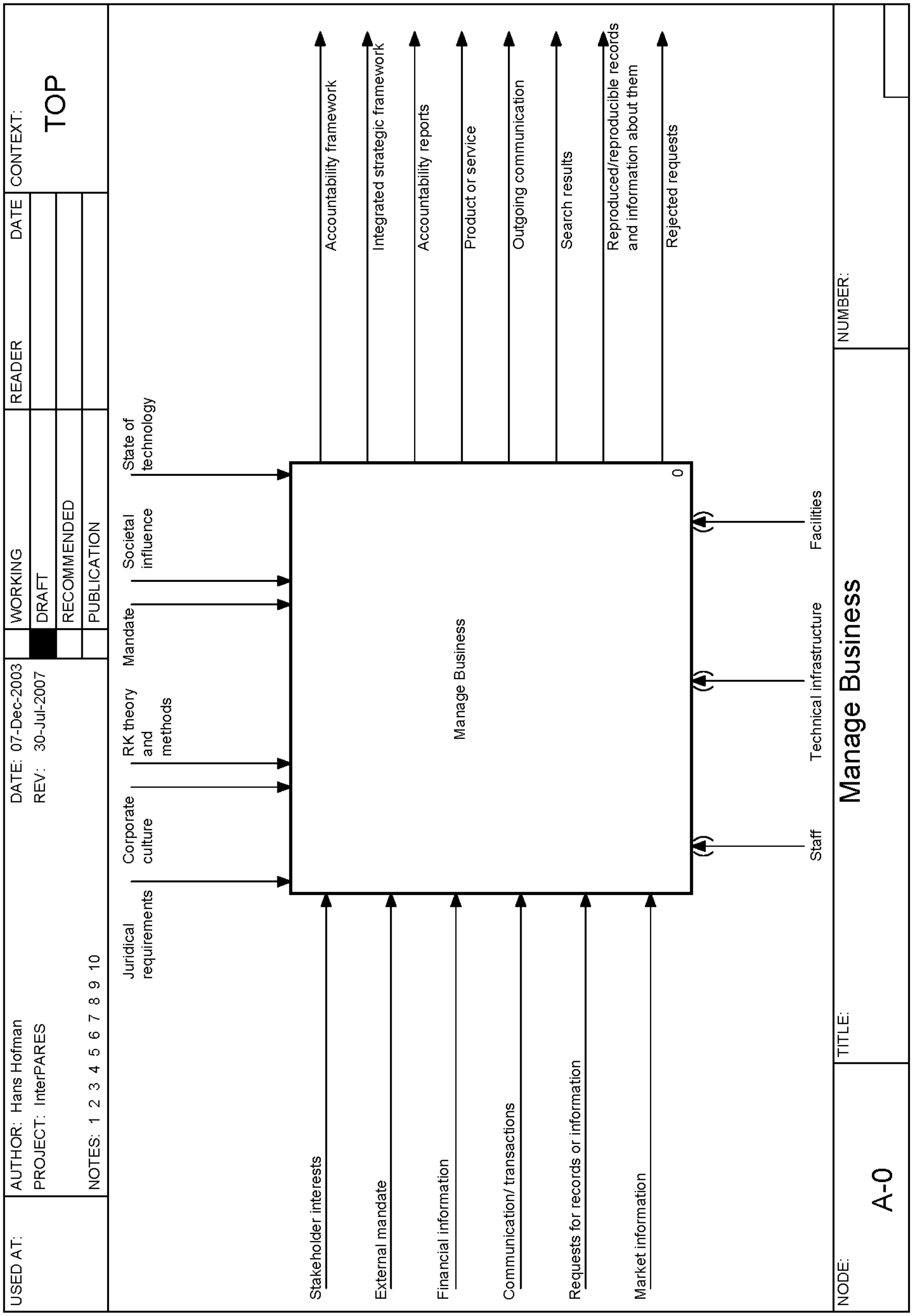
APPENDIX 15

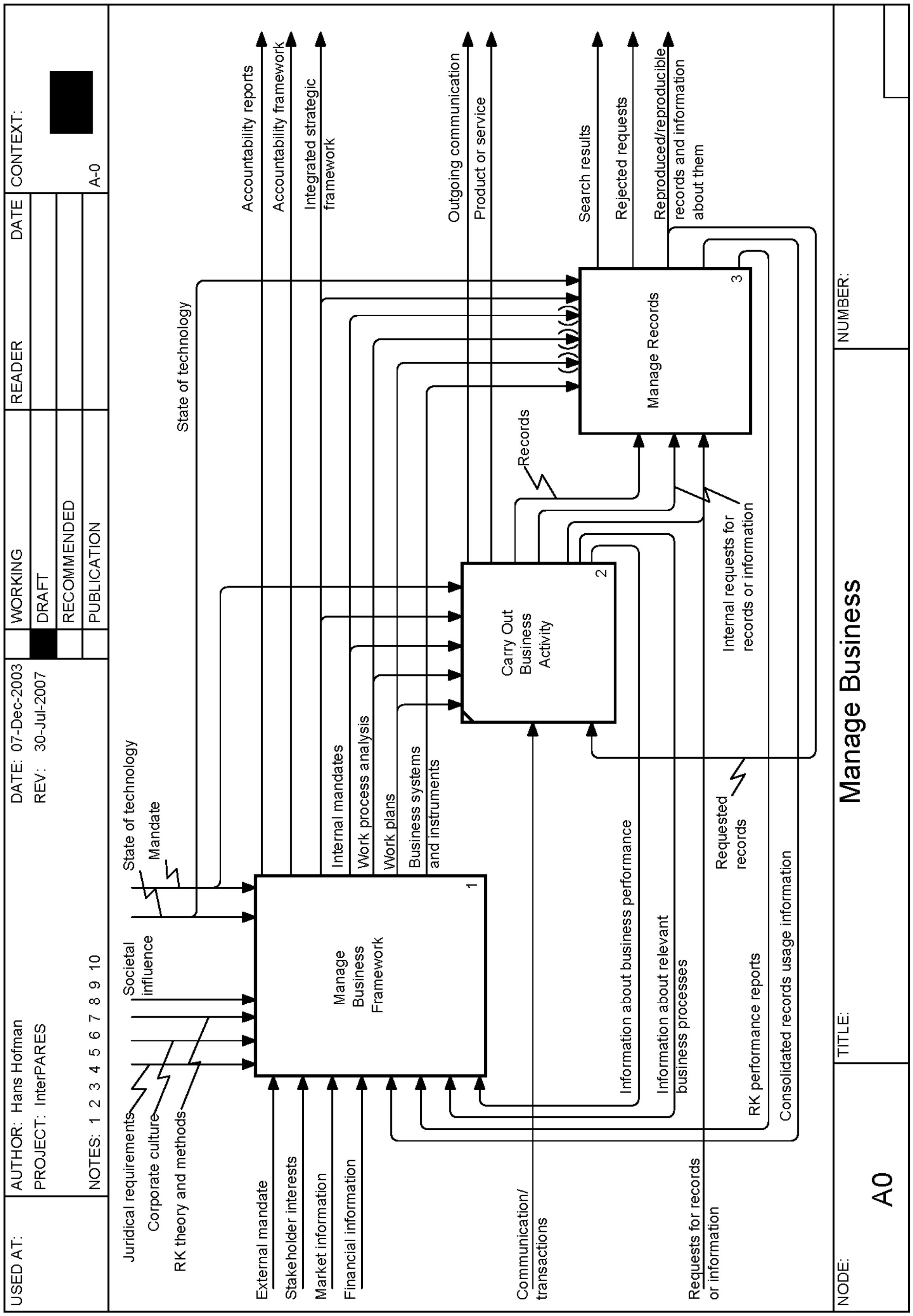
Business-driven Recordkeeping Model

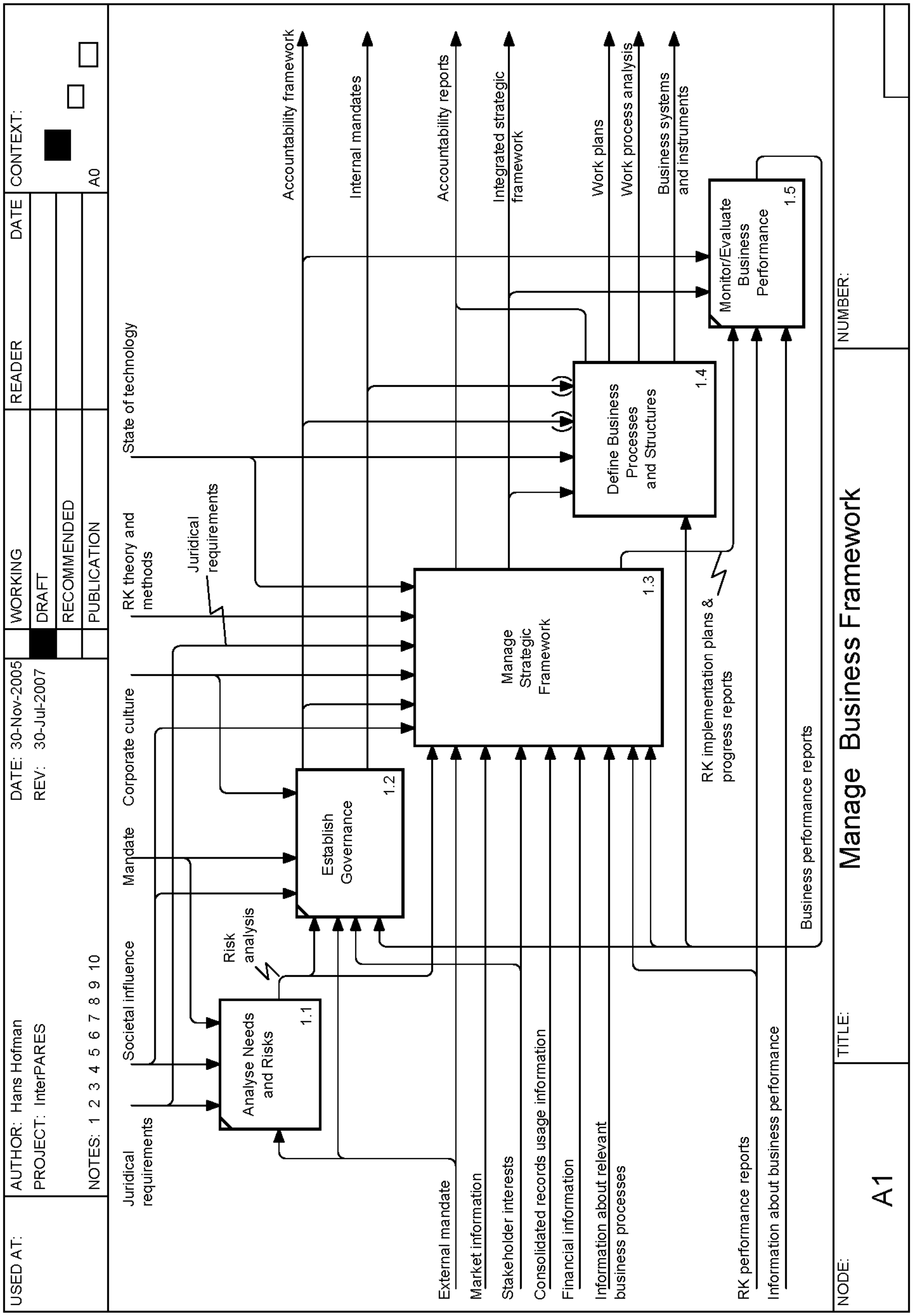
Diagrams and Definitions

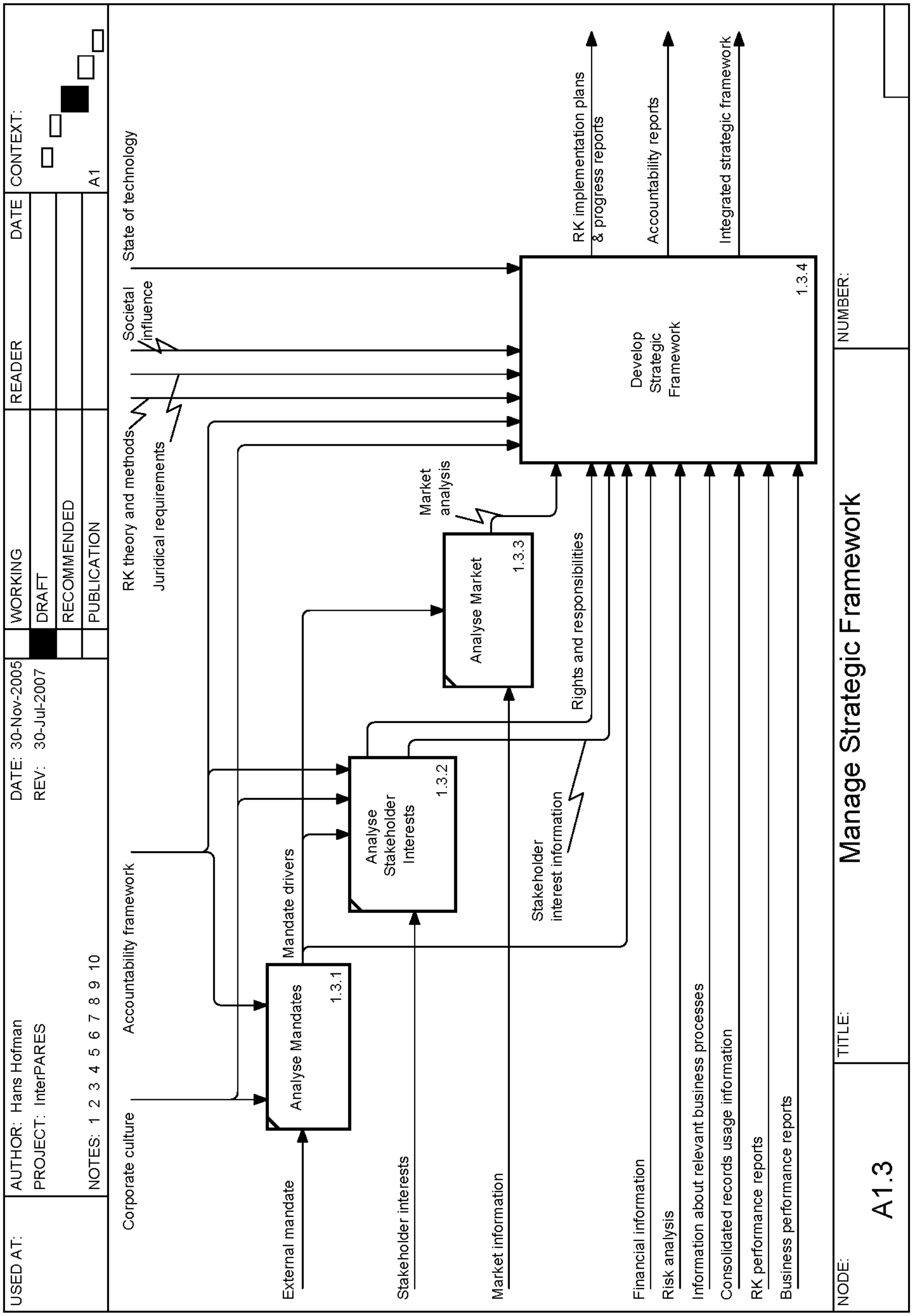
by

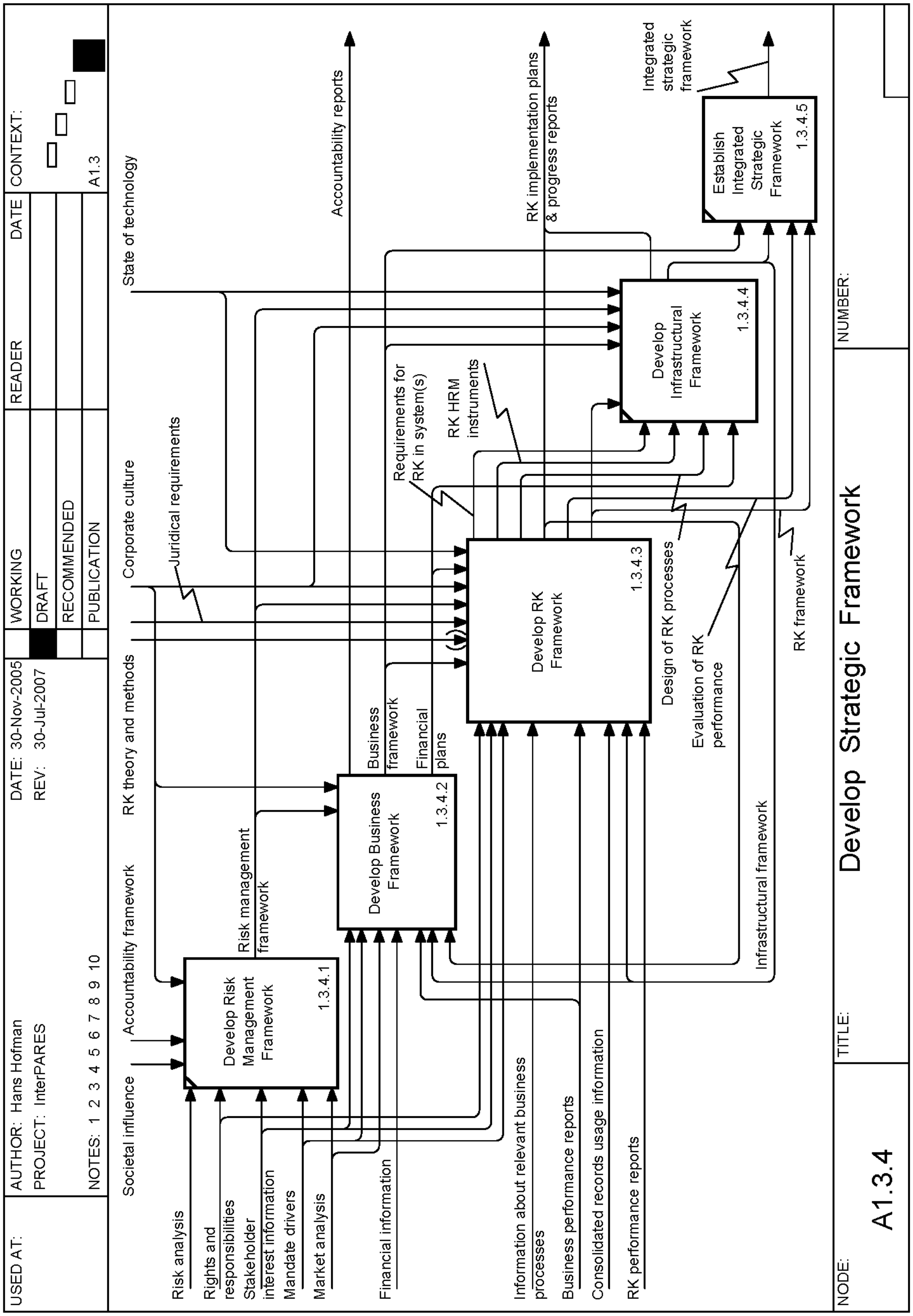
Hans Hofman
National Archives of the Netherlands

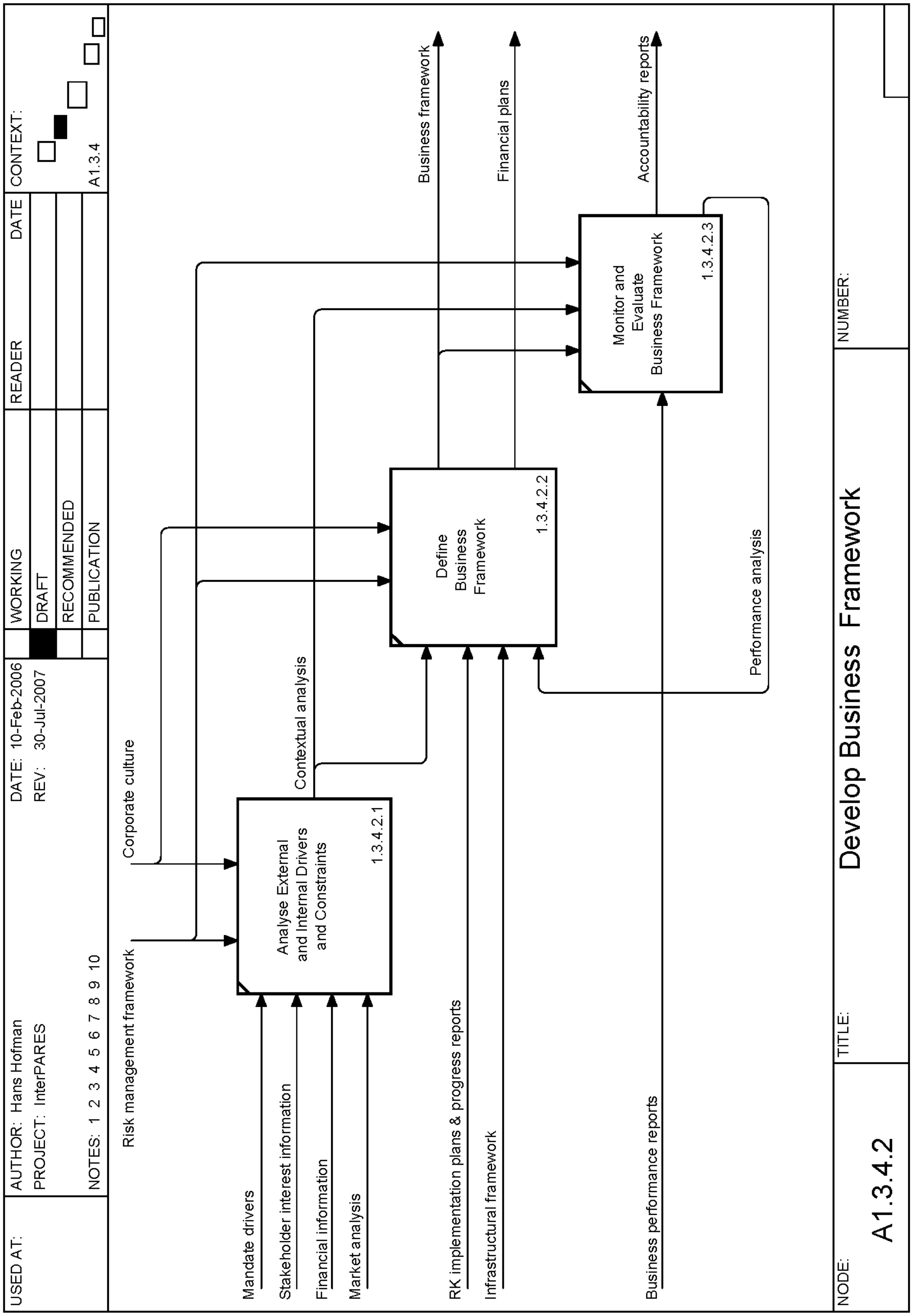

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 07-Dec-2003
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
TOP
Juridical requirements
Corporate culture
RK theory and methods
Mandate
Societal influence
State of technology
Stakeholder interests
External mandate
Financial information
Communication/ transactions
Requests for records or information
Market information
Manage Business
0
Accountability framework
Integrated strategic framework
Accountability reports
Product or service
Outgoing communication
Search results
Reproduced/reproducible records and information about them
Rejected requests
Staff
Technical infrastructure
Facilities
NODE:
A-0
TITLE:
Manage Business
NUMBER:

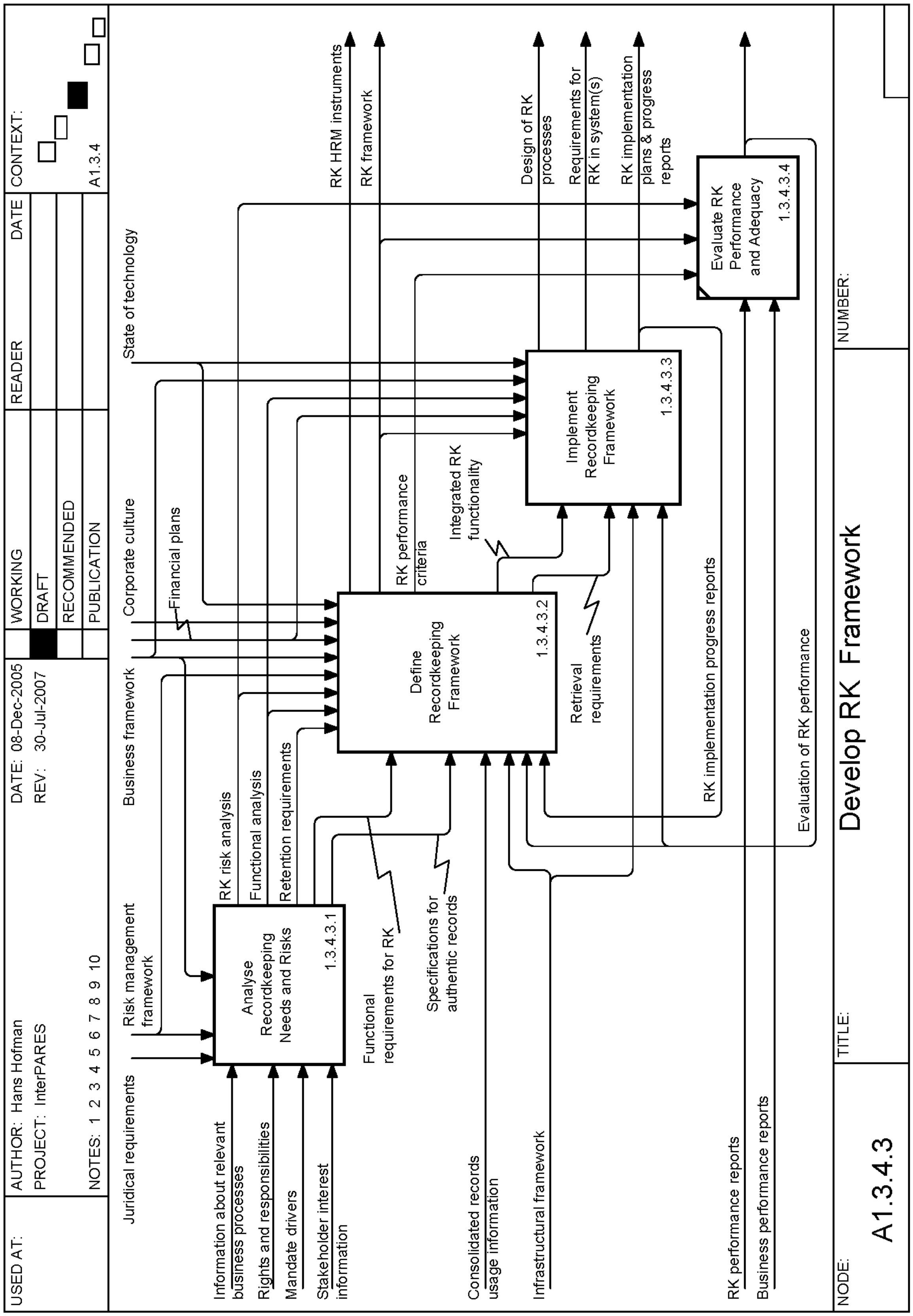

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 07-Dec-2003
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A-0
Juridical requirements
Corporate culture
RK theory and methods
Societal influence
State of technology
Mandate
External mandate
Stakeholder interests
Market information
Financial information
Manage Business Framework
1
Internal mandates
Work process analysis
Work plans
Business systems and instruments
Accountability reports
Accountability framework
Integrated strategic framework
State of technology
Communication/ transactions
Carry Out Business Activity
2
Outgoing communication
Product or service
Records
Internal requests for records or information
Requested records
Manage Records
3
Search results
Rejected requests
Reproduced/reproducible records and information about them
Information about business performance
Information about relevant business processes
RK performance reports
Consolidated records usage information
Requests for records or information
NODE:
A0
TITLE:
Manage Business
NUMBER:

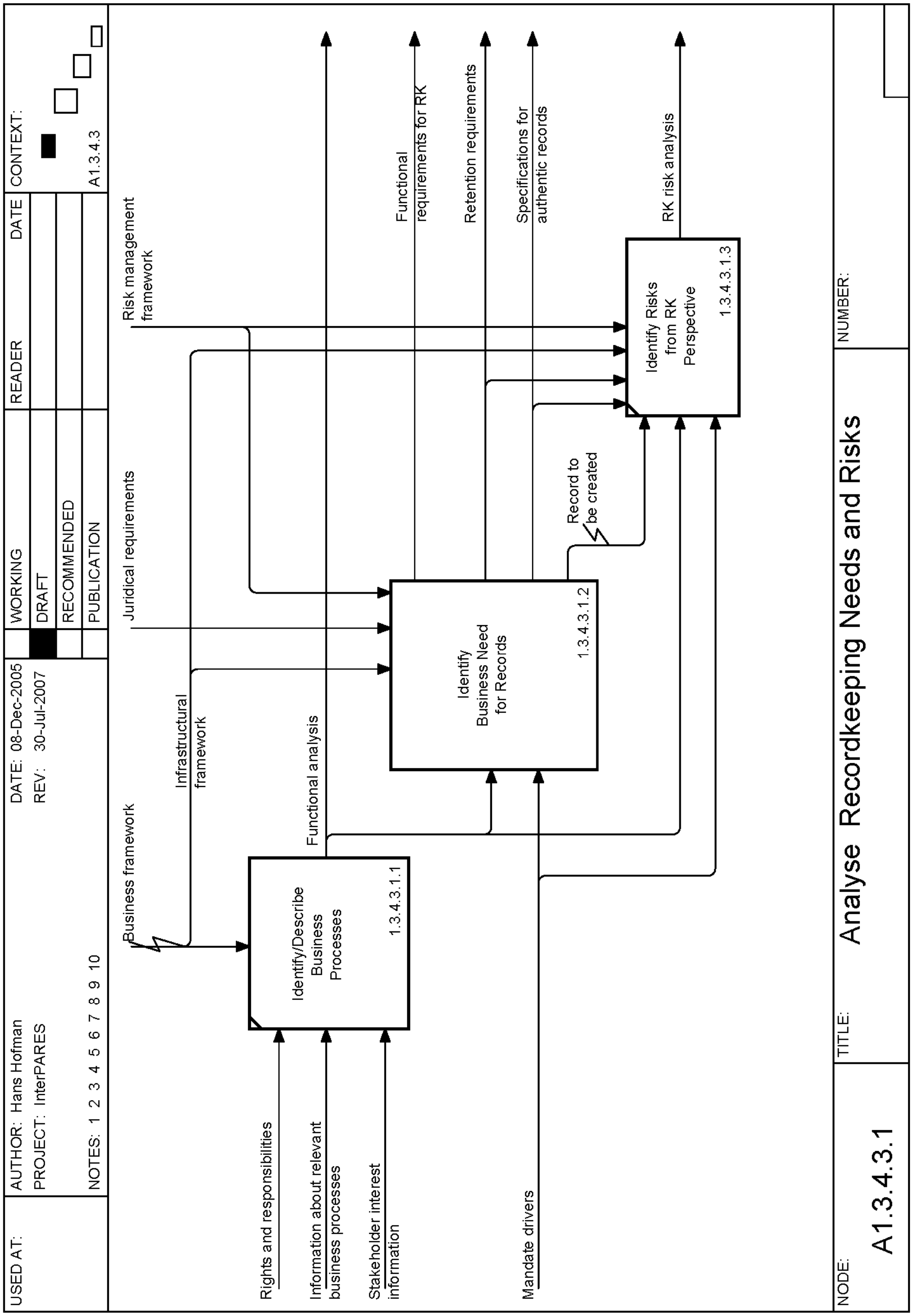

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
DATE: 30-Nov-2005
REV: 30-Jul-2007
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A0
Juridical requirements
Societal influence
Mandate
Corporate culture
RK theory and methods
Juridical requirements
State of technology
Risk analysis
Analyse Needs and Risks
1.1
Establish Governance
1.2
Manage Strategic Framework
1.3
Define Business Processes and Structures
1.4
Monitor/Evaluate Business Performance
1.5
Accountability framework
Internal mandates
Accountability reports
Integrated strategic framework
Work plans
Work process analysis
Business systems and instruments
RK implementation plans & progress reports
Business performance reports
External mandate
Market information
Stakeholder interests
Consolidated records usage information
Financial information
Information about relevant business processes
RK performance reports
Information about business performance
NODE:
A1
TITLE:
Manage Business Framework
NUMBER:

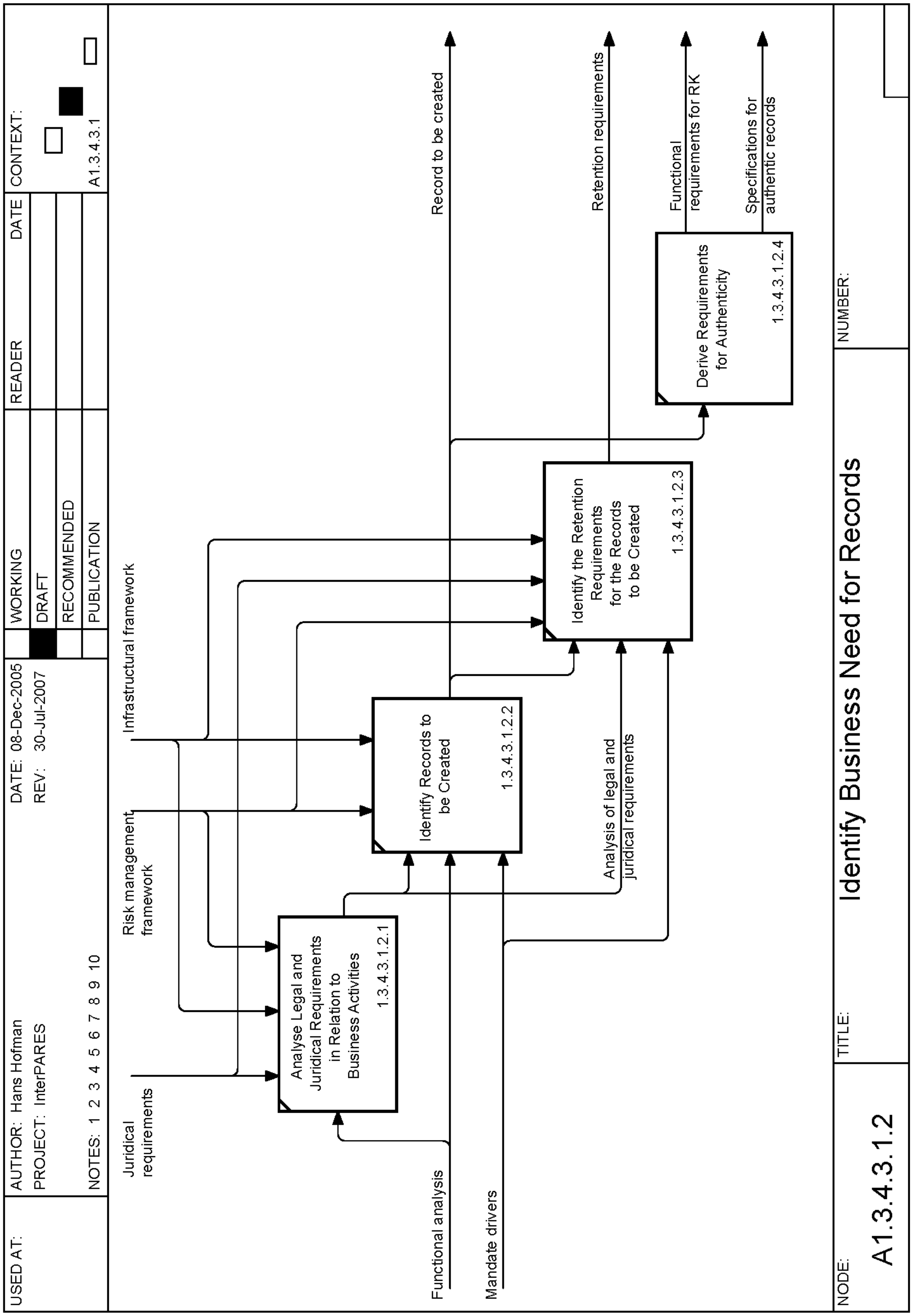

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 30-Nov-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1
Corporate culture
Accountability framework
RK theory and methods
Juridical requirements
Societal influence
State of technology
External mandate
Stakeholder interests
Market information
Analyse Mandates
1.3.1
Mandate drivers
Analyse Stakeholder Interests
1.3.2
Stakeholder interest information
Analyse Market
1.3.3
Market analysis
Rights and responsibilities
Develop Strategic Framework
1.3.4
RK implementation plans & progress reports
Accountability reports
Integrated strategic framework
Financial information
Risk analysis
Information about relevant business processes
Consolidated records usage information
RK performance reports
Business performance reports
NODE:
A1.3
TITLE:
Manage Strategic Framework
NUMBER:

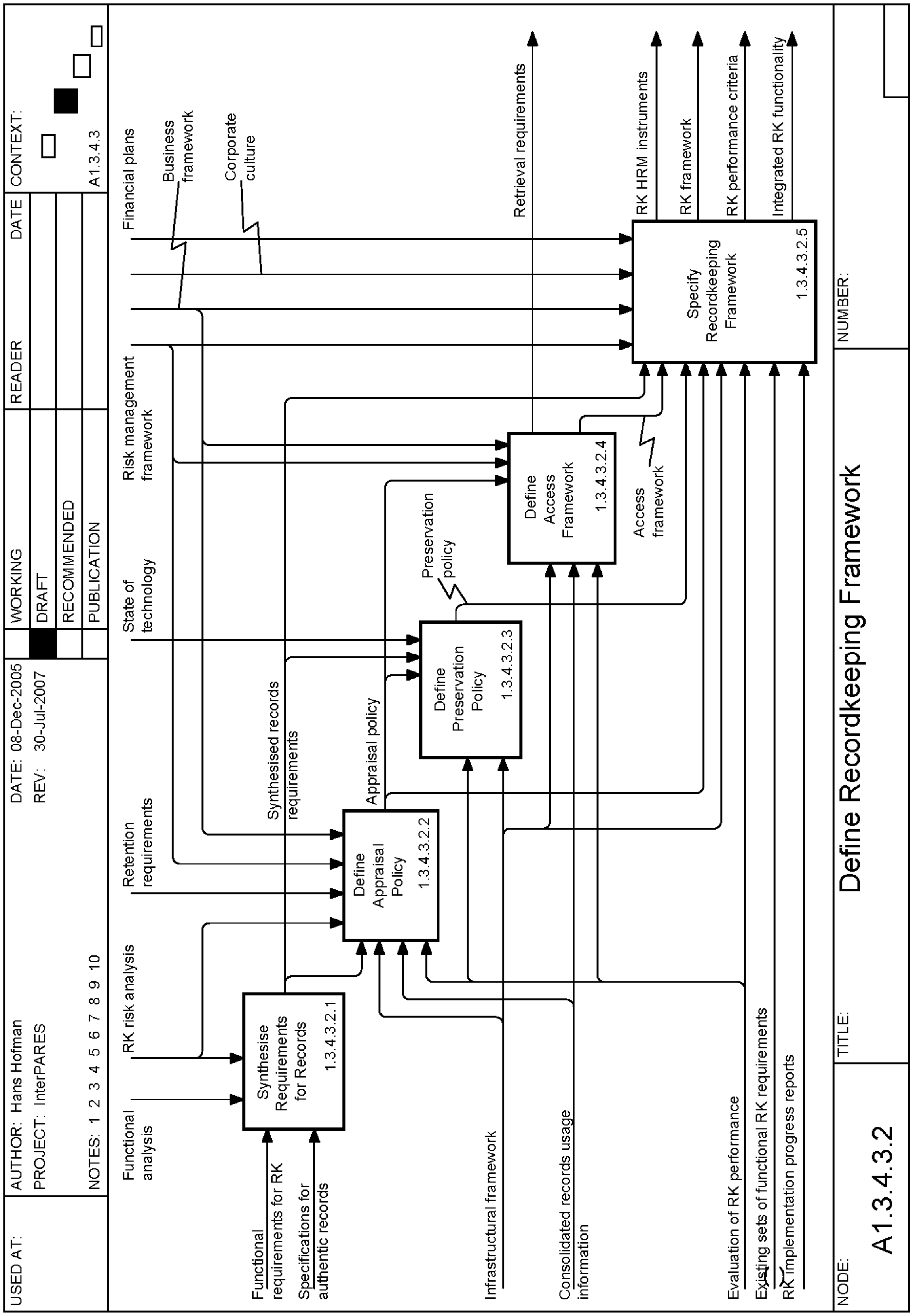

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 30-Nov-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3
Societal influence
Accountability framework
RK theory and methods
Corporate culture
Juridical requirements
State of technology
Risk analysis
Rights and responsibilities
Stakeholder interest information
Mandate drivers
Market analysis
Financial information
Information about relevant business processes
Business performance reports
Consolidated records usage information
RK performance reports
Develop Risk Management Framework
1.3.4.1
Risk management framework
Develop Business Framework
1.3.4.2
Business framework
Financial plans
Develop RK Framework
1.3.4.3
Accountability reports
Requirements for RK in system(s)
RK HRM instruments
Design of RK processes
Evaluation of RK performance
RK framework
Infrastructural framework
Develop Infrastructural Framework
1.3.4.4
RK implementation plans & progress reports
Establish Integrated Strategic Framework
1.3.4.5
Integrated strategic framework
NODE:
A1.3.4
TITLE:
Develop Strategic Framework
NUMBER:


USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 10-Feb-2006
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4
Risk management framework
Corporate culture
Mandate drivers
Stakeholder interest information
Financial information
Market analysis
Analyse External and Internal Drivers and Constraints
1.3.4.2.1
Contextual analysis
RK implementation plans & progress reports
Infrastructural framework
Define Business Framework
1.3.4.2.2
Business framework
Financial plans
Business performance reports
Monitor and Evaluate Business Framework
1.3.4.2.3
Accountability reports
Performance analysis
NODE:
A1.3.4.2
TITLE:
Develop Business Framework
NUMBER:

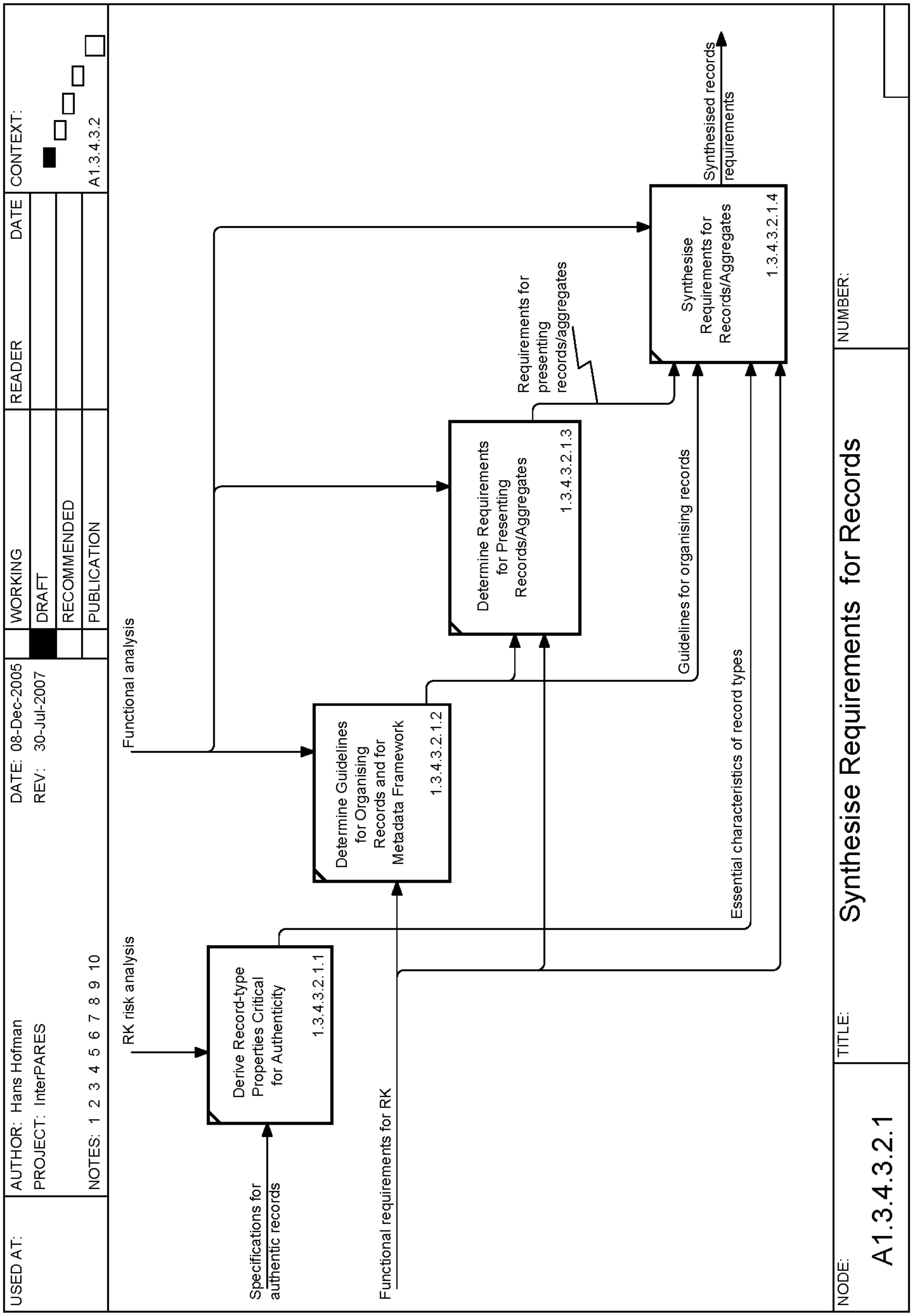

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4
Juridical requirements
Risk management framework
Information about relevant business processes
Rights and responsibilities
Mandate drivers
Stakeholder interest information
Analyse Recordkeeping Needs and Risks
1.3.4.3.1
RK risk analysis
Functional analysis
Retention requirements
Functional requirements for RK
Specifications for authentic records
Business framework
Corporate culture
Financial plans
State of technology
Define Recordkeeping Framework
1.3.4.3.2
RK performance criteria
Integrated RK functionality
Retrieval requirements
Consolidated records usage information
Infrastructural framework
Implement Recordkeeping Framework
1.3.4.3.3
RK HRM instruments
RK framework
Design of RK processes
Requirements for RK in system(s)
RK implementation plans & progress reports
RK implementation progress reports
Evaluation of RK performance
Evaluate RK Performance and Adequacy
1.3.4.3.4
RK performance reports
Business performance reports
NODE:
A1.3.4.3
TITLE:
Develop RK Framework
NUMBER:

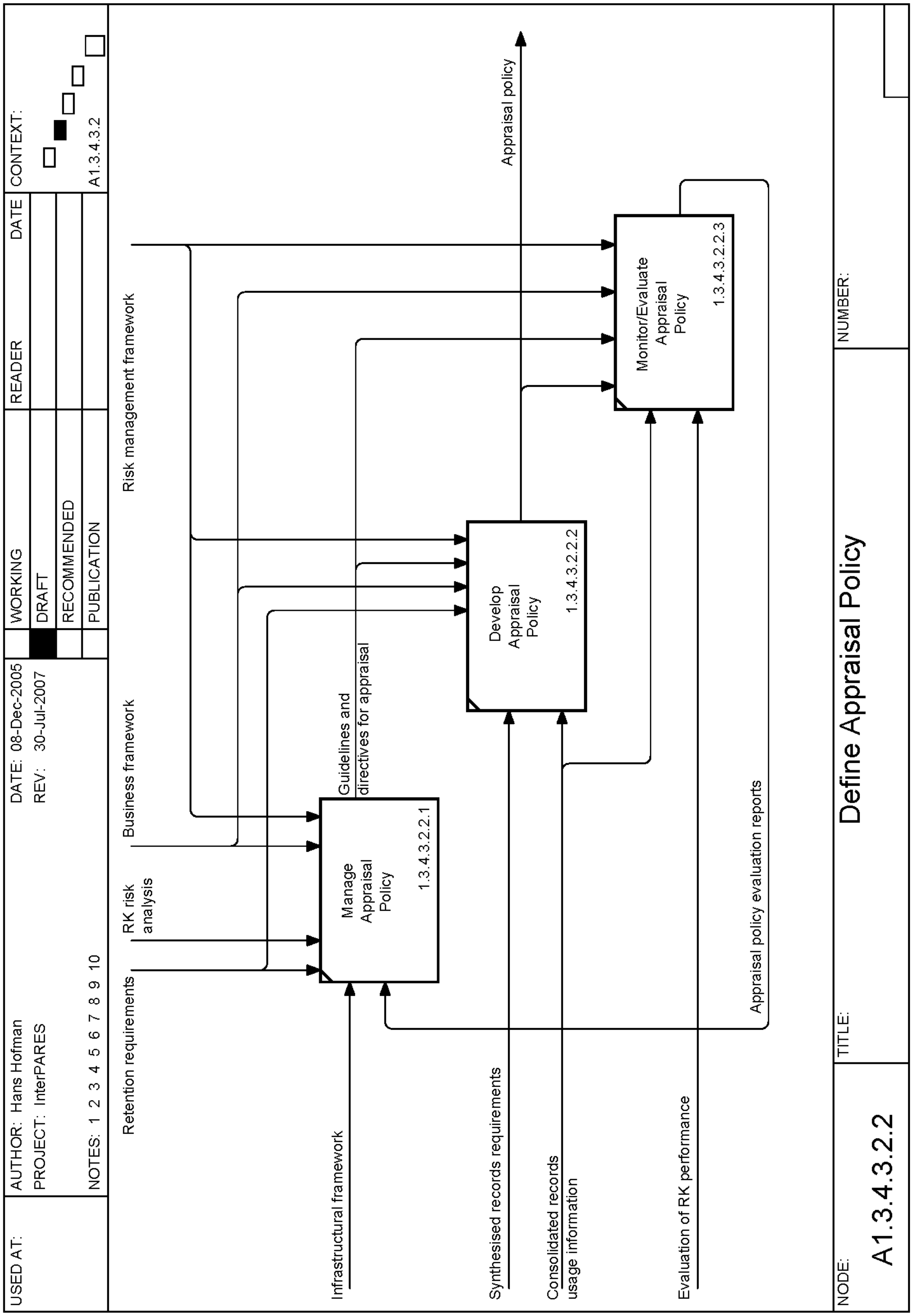

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3
Rights and responsibilities
Information about relevant business processes
Stakeholder interest information
Mandate drivers
Business framework
Infrastructural framework
Functional analysis
Juridical requirements
Risk management framework
Identify/Describe Business Processes
1.3.4.3.1.1
Identify Business Need for Records
1.3.4.3.1.2
Record to be created
Identify Risks from RK Perspective
1.3.4.3.1.3
Functional requirements for RK
Retention requirements
Specifications for authentic records
RK risk analysis
NODE:
A1.3.4.3.1
TITLE:
Analyse Recordkeeping Needs and Risks
NUMBER:

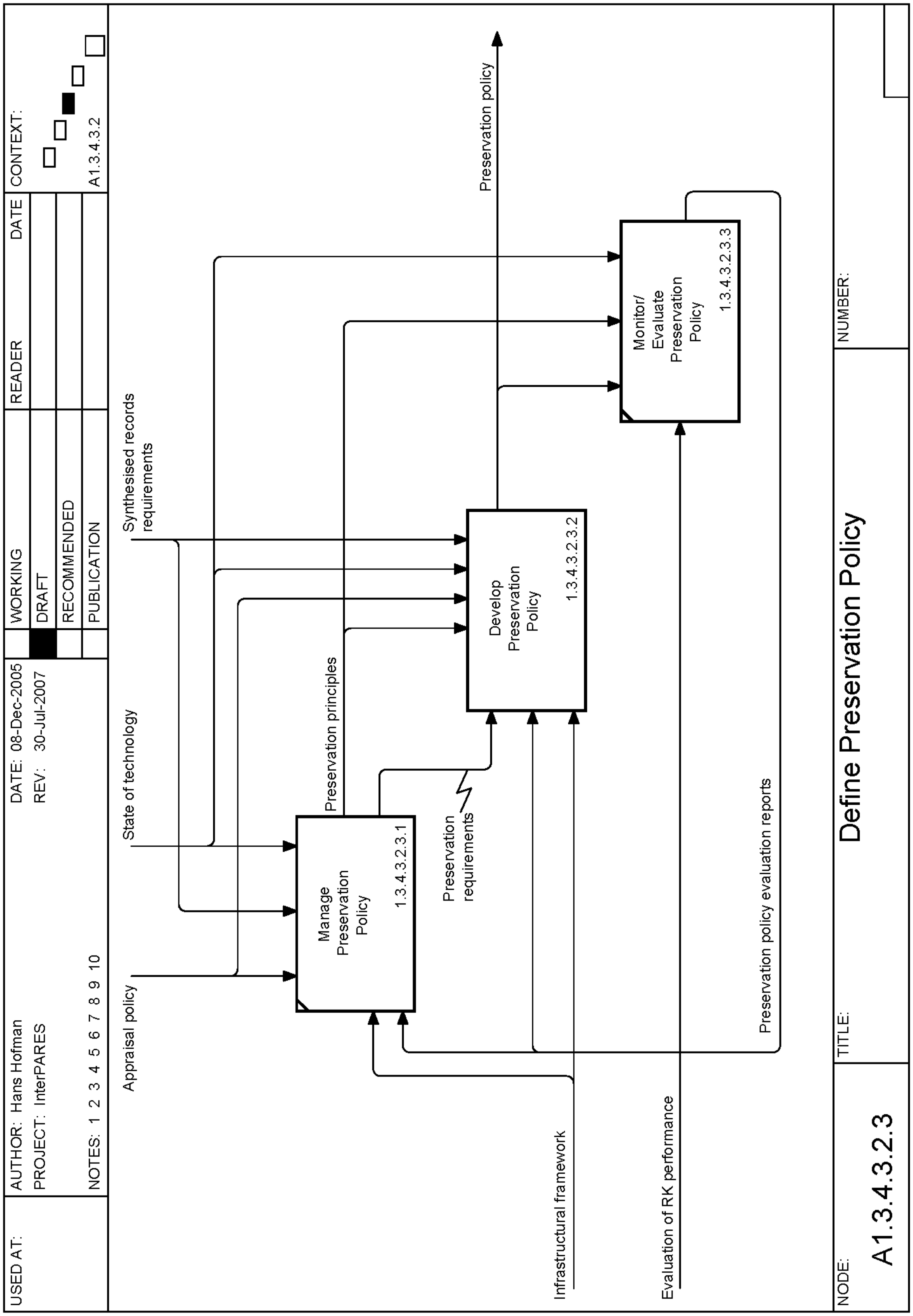

USED AT:
AUTHOR: Hans Hofman
DATE: 08-Dec-2005
PROJECT: InterPARES
REV: 30-Jul-2007
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.1
Juridical requirements
Risk management framework
Infrastructural framework
Functional analysis
Mandate drivers
Analysis of legal and juridical requirements
Analyse Legal and Juridical Requirements in Relation to Business Activities
1.3.4.3.1.2.1
Identify Records to be Created
1.3.4.3.1.2.2
Identify the Retention Requirements for the Records to be Created
1.3.4.3.1.2.3
Derive Requirements for Authenticity
1.3.4.3.1.2.4
Record to be created
Retention requirements
Functional requirements for RK
Specifications for authentic records
NODE:
A1.3.4.3.1.2
TITLE:
Identify Business Need for Records
NUMBER:

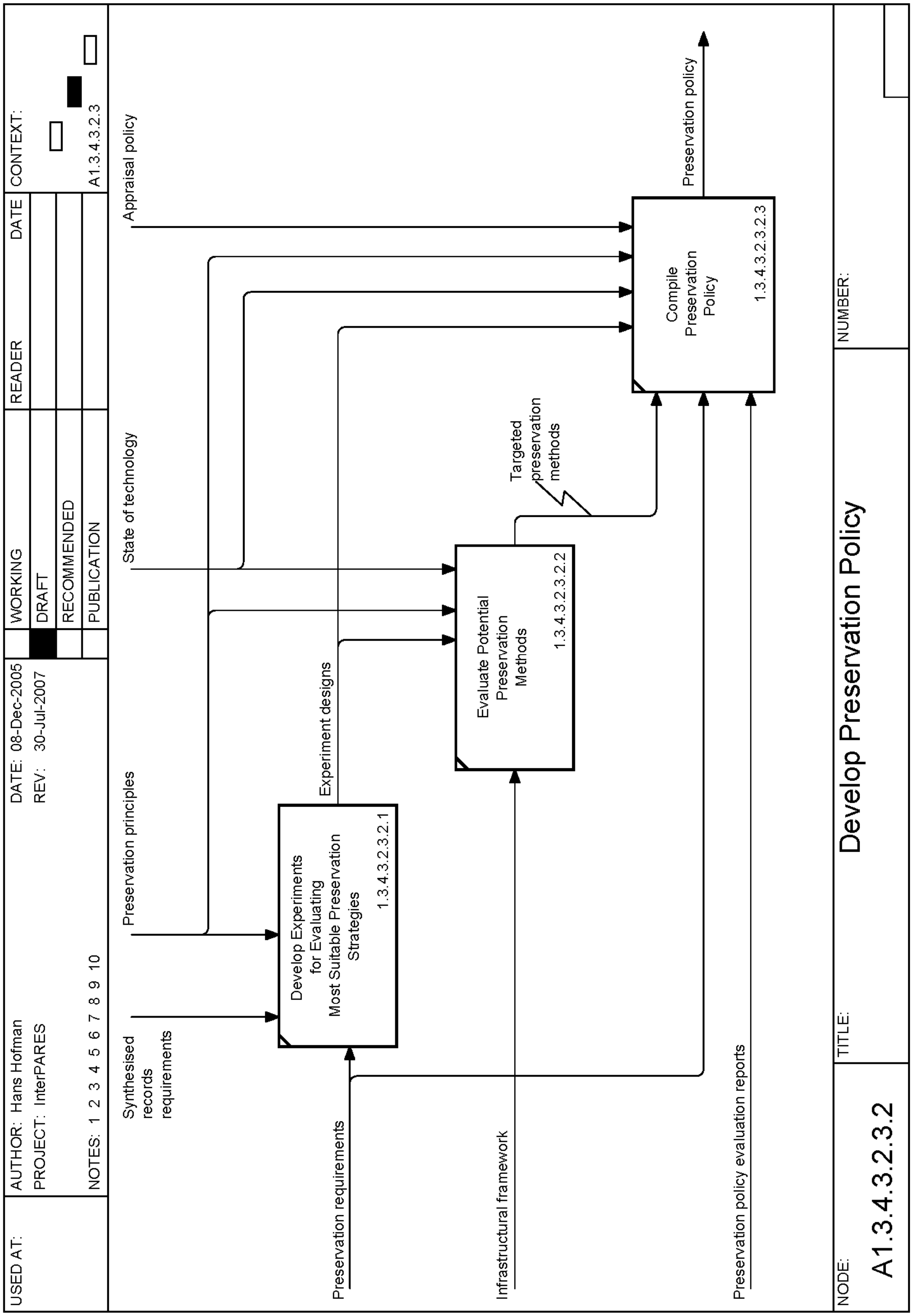


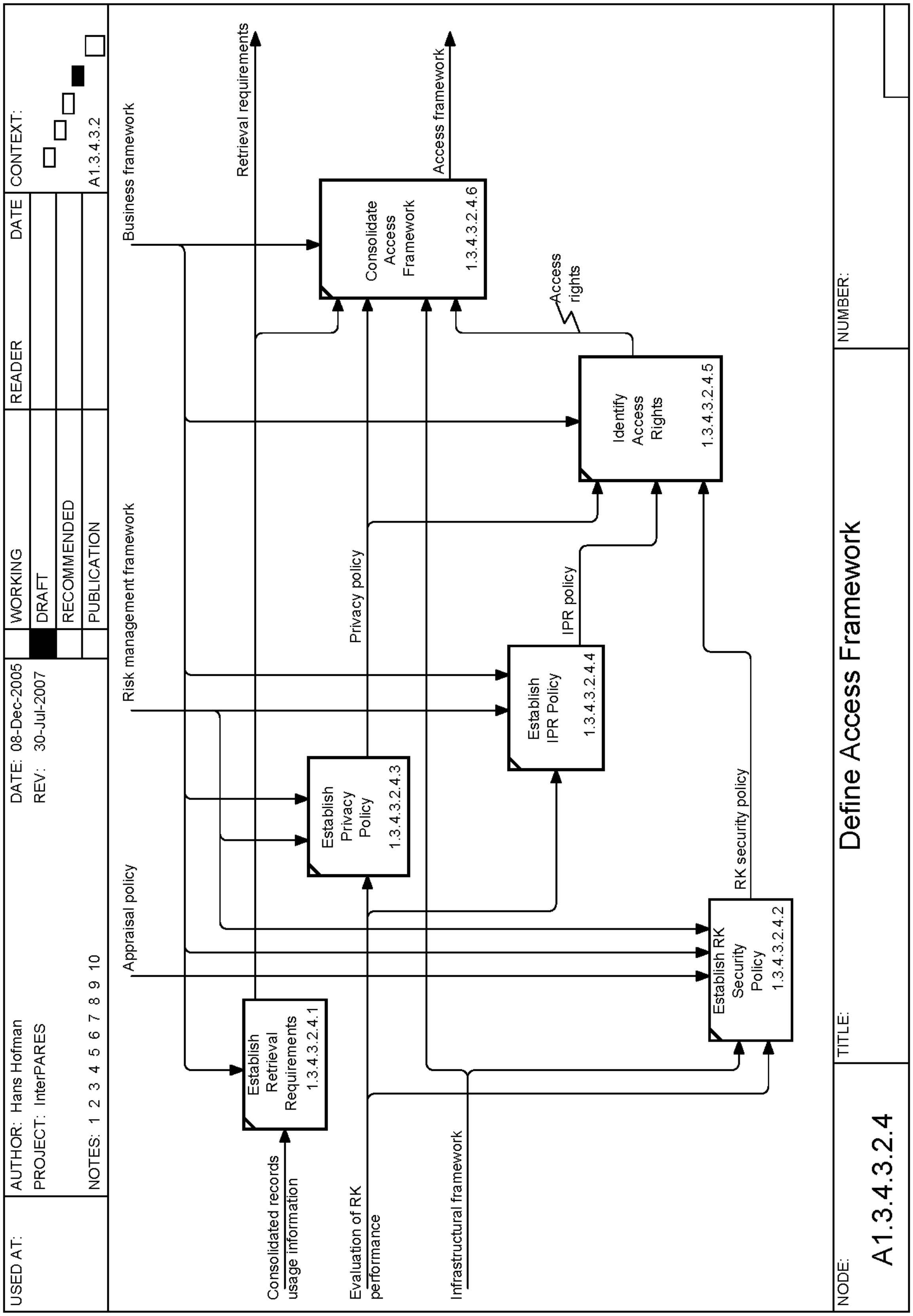
USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3
Functional analysis
RK risk analysis
Retention requirements
State of technology
Risk management framework
Financial plans
Business framework
Corporate culture
Retrieval requirements
RK HRM instruments
RK framework
RK performance criteria
Integrated RK functionality
Synthesise Requirements for Records
1.3.4.3.2.1
Define Appraisal Policy
1.3.4.3.2.2
Define Preservation Policy
1.3.4.3.2.3
Define Access Framework
1.3.4.3.2.4
Specify Recordkeeping Framework
1.3.4.3.2.5
Synthesised records requirements
Appraisal policy
Preservation policy
Access framework
Functional requirements for RK
Specifications for authentic records
Infrastructural framework
Consolidated records usage information
Evaluation of RK performance
Existing sets of functional RK requirements
RK implementation progress reports
NODE:
A1.3.4.3.2
TITLE:
Define Recordkeeping Framework
NUMBER:

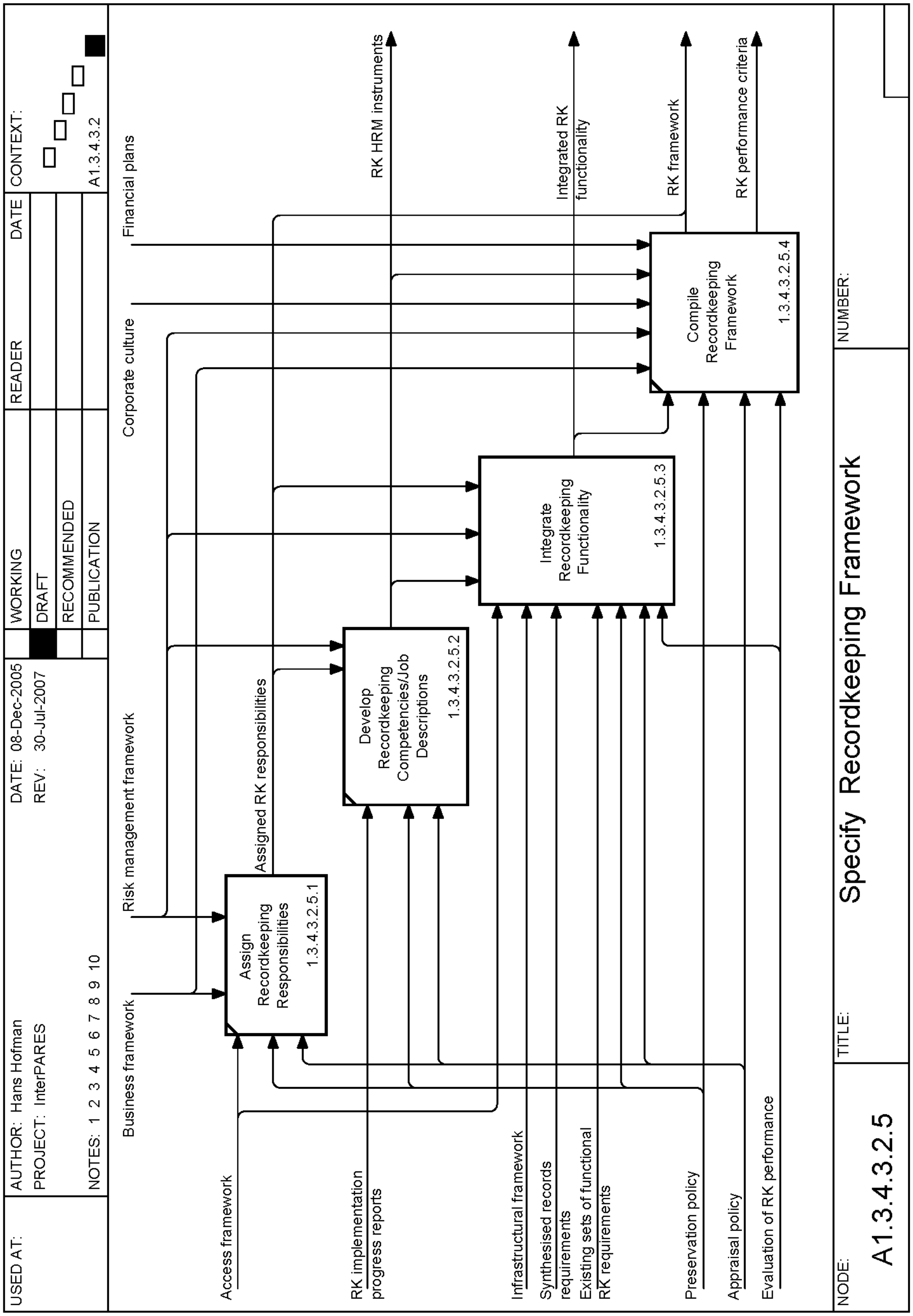

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.2
RK risk analysis
Specifications for authentic records
Derive Record-type Properties Critical for Authenticity
1.3.4.3.2.1.1
Functional requirements for RK
Functional analysis
Determine Guidelines for Organising Records and for Metadata Framework
1.3.4.3.2.1.2
Essential characteristics of record types
Guidelines for organising records
Determine Requirements for Presenting Records/Aggregates
1.3.4.3.2.1.3
Requirements for presenting records/aggregates
Synthesise Requirements for Records/Aggregates
1.3.4.3.2.1.4
Synthesised records requirements
NODE:
A1.3.4.3.2.1
TITLE:
Synthesise Requirements for Records
NUMBER:

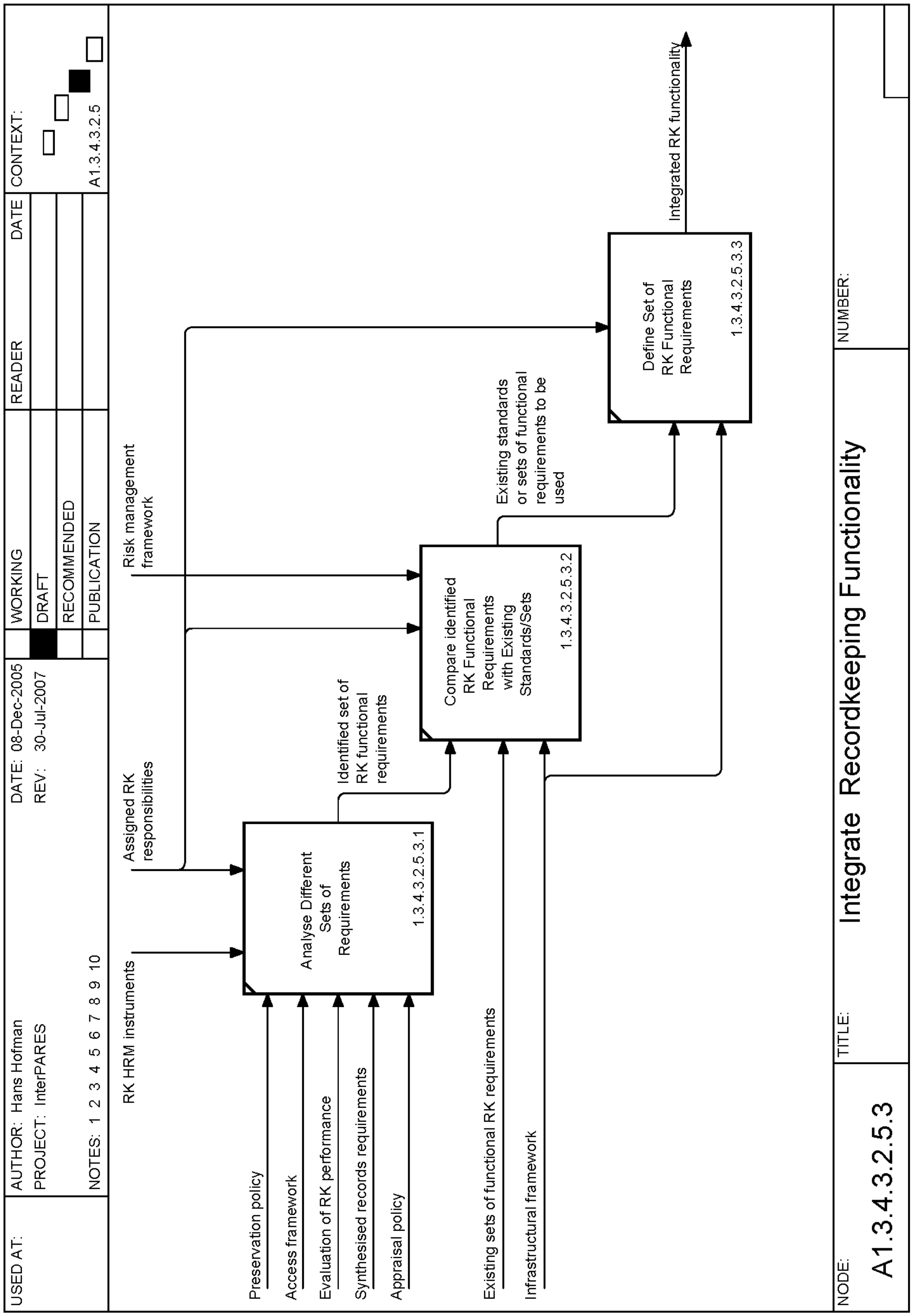

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.2
Retention requirements
RK risk analysis
Business framework
Risk management framework
Infrastructural framework
Manage Appraisal Policy
1.3.4.3.2.2.1
Guidelines and directives for appraisal
Develop Appraisal Policy
1.3.4.3.2.2.2
Appraisal policy
Monitor/Evaluate Appraisal Policy
1.3.4.3.2.2.3
Appraisal policy evaluation reports
Synthesised records requirements
Consolidated records usage information
Evaluation of RK performance
NODE:
A1.3.4.3.2.2
TITLE:
Define Appraisal Policy
NUMBER:

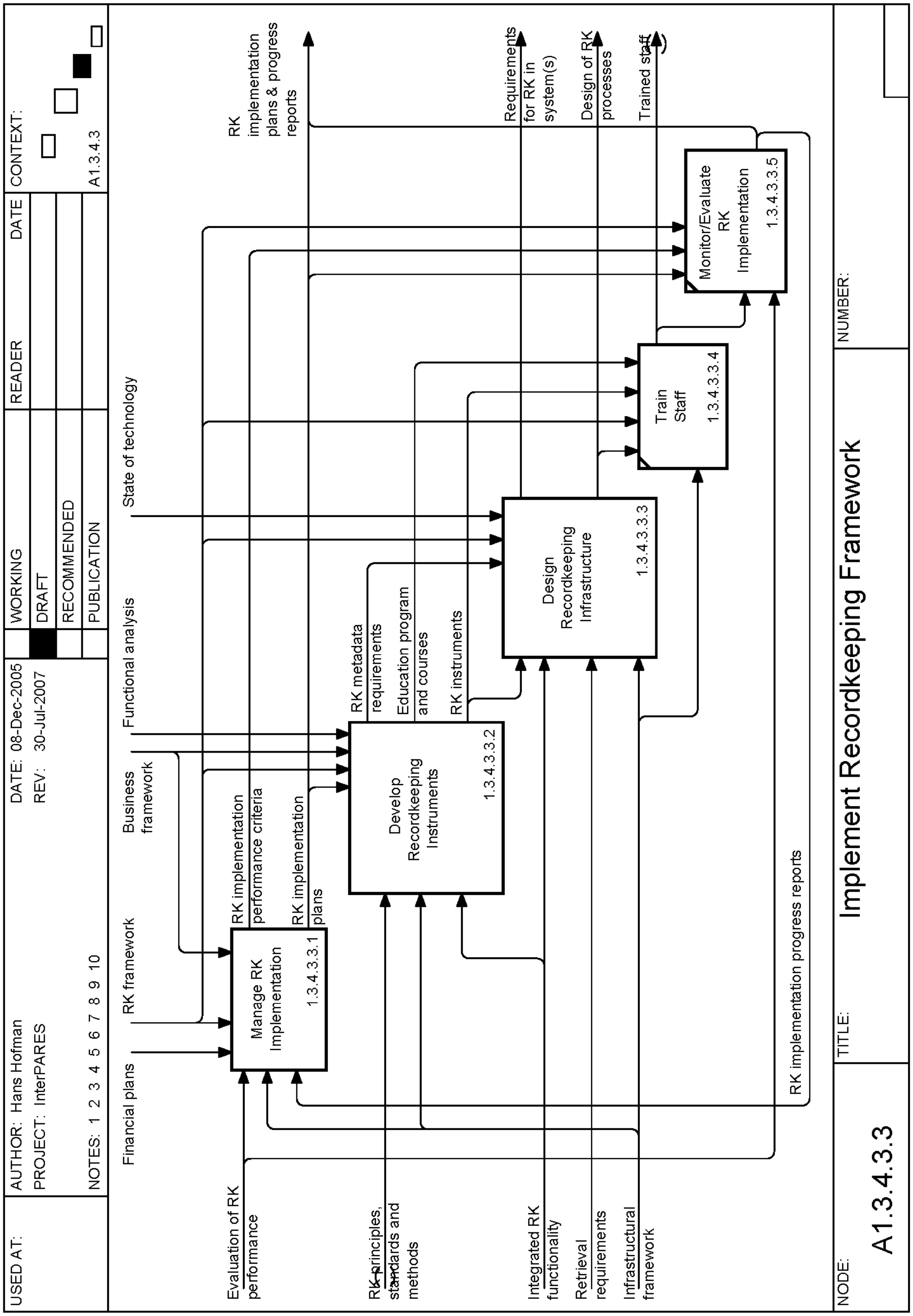

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.2
Appraisal policy
State of technology
Synthesised records requirements
Preservation principles
Manage Preservation Policy
1.3.4.3.2.3.1
Preservation requirements
Develop Preservation Policy
1.3.4.3.2.3.2
Preservation policy
Monitor/ Evaluate Preservation Policy
1.3.4.3.2.3.3
Preservation policy evaluation reports
Infrastructural framework
Evaluation of RK performance
NODE:
A1.3.4.3.2.3
TITLE:
Define Preservation Policy
NUMBER:

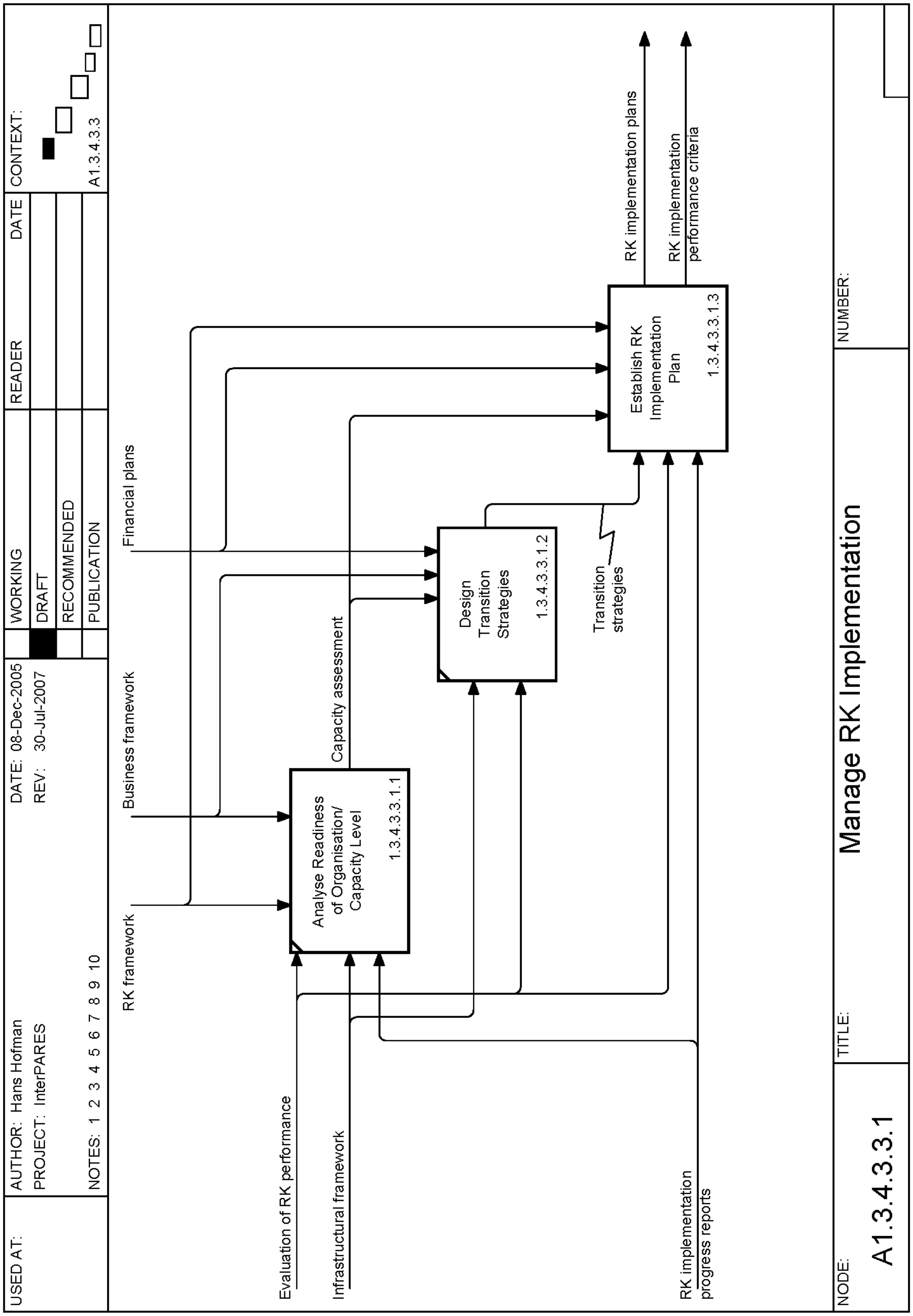

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.2.3
Synthesised records requirements
Preservation principles
State of technology
Appraisal policy
Preservation requirements
Infrastructural framework
Preservation policy evaluation reports
Develop Experiments for Evaluating Most Suitable Preservation Strategies
1.3.4.3.2.3.2.1
Experiment designs
Evaluate Potential Preservation Methods
1.3.4.3.2.3.2.2
Targeted preservation methods
Compile Preservation Policy
1.3.4.3.2.3.2.3
Preservation policy
NODE:
A1.3.4.3.2.3.2
TITLE:
Develop Preservation Policy
NUMBER:

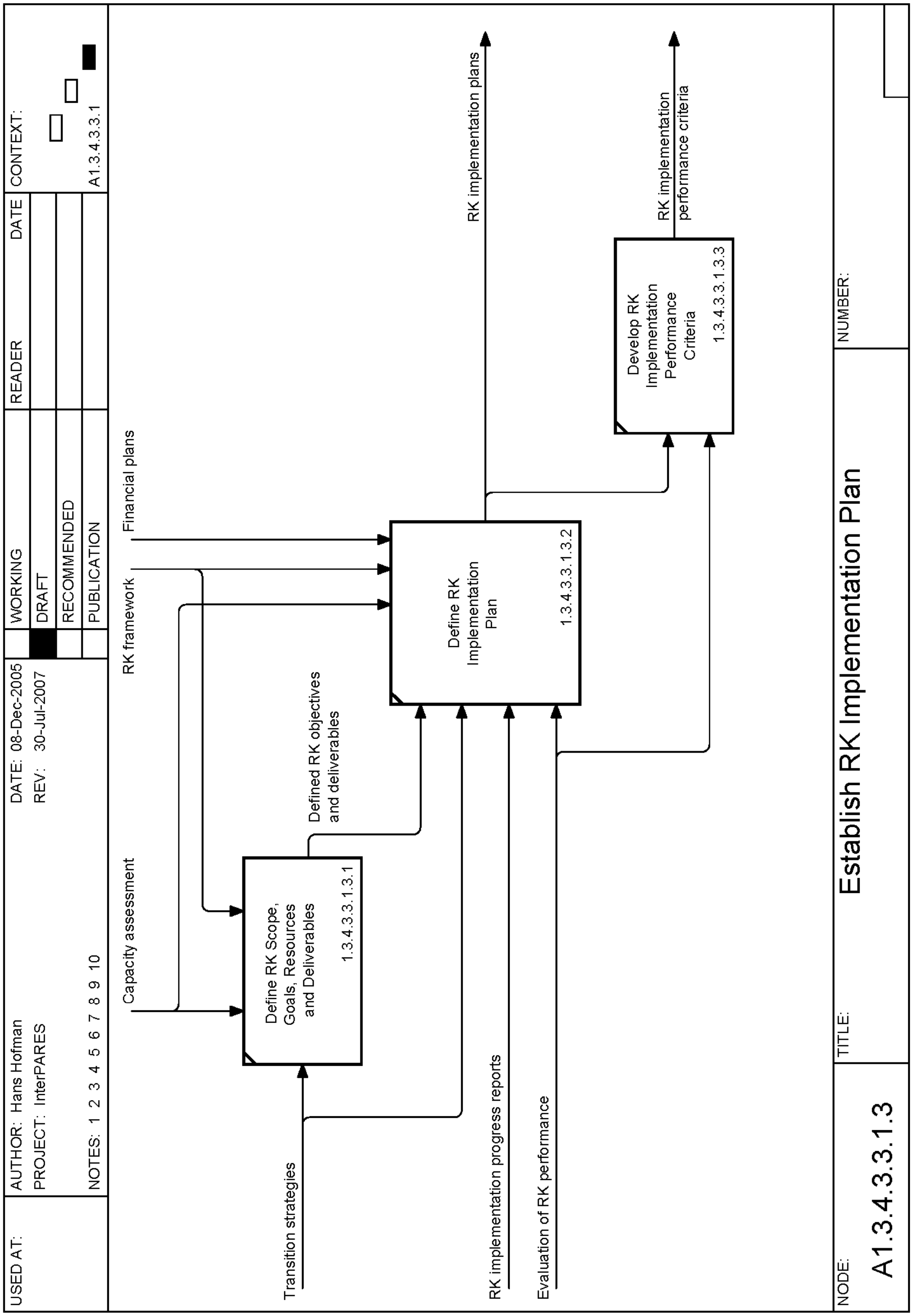

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.2
Consolidated records usage information
Evaluation of RK performance
Infrastructural framework
Appraisal policy
Risk management framework
Business framework
Retrieval requirements
Access framework
Privacy policy
IPR policy
RK security policy
Access rights
Establish Retrieval Requirements
1.3.4.3.2.4.1
Establish RK Security Policy
1.3.4.3.2.4.2
Establish Privacy Policy
1.3.4.3.2.4.3
Establish IPR Policy
1.3.4.3.2.4.4
Identify Access Rights
1.3.4.3.2.4.5
Consolidate Access Framework
1.3.4.3.2.4.6
NODE:
A1.3.4.3.2.4
TITLE:
Define Access Framework
NUMBER:

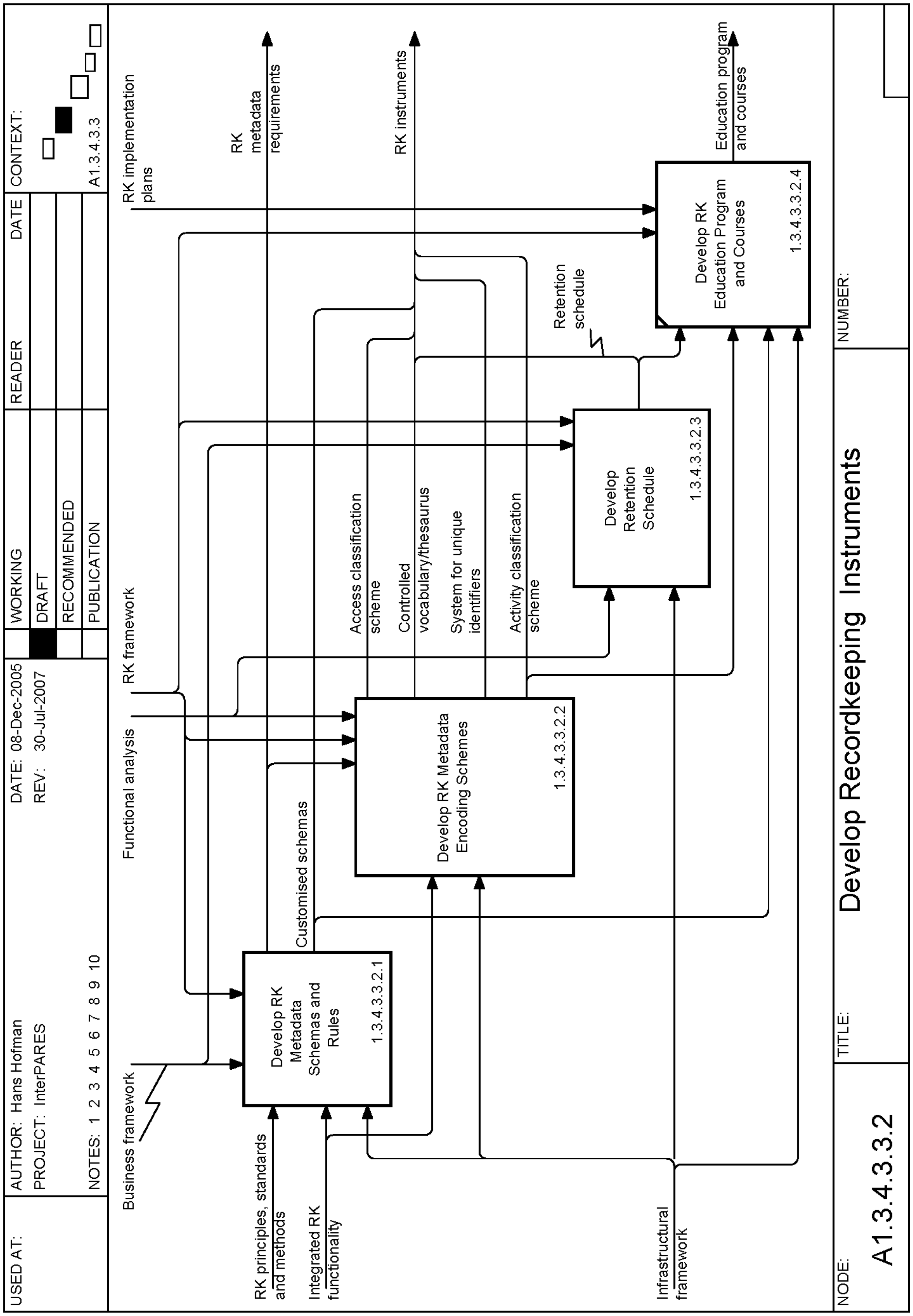

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.2
Business framework
Risk management framework
Corporate culture
Financial plans
Assigned RK responsibilities
RK HRM instruments
Integrated RK functionality
RK framework
RK performance criteria
Assign Recordkeeping Responsibilities
1.3.4.3.2.5.1
Develop Recordkeeping Competencies/Job Descriptions
1.3.4.3.2.5.2
Integrate Recordkeeping Functionality
1.3.4.3.2.5.3
Compile Recordkeeping Framework
1.3.4.3.2.5.4
Access framework
RK implementation progress reports
Infrastructural framework
Synthesised records requirements
Existing sets of functional RK requirements
Preservation policy
Appraisal policy
Evaluation of RK performance
NODE:
A1.3.4.3.2.5
TITLE:
Specify Recordkeeping Framework
NUMBER:

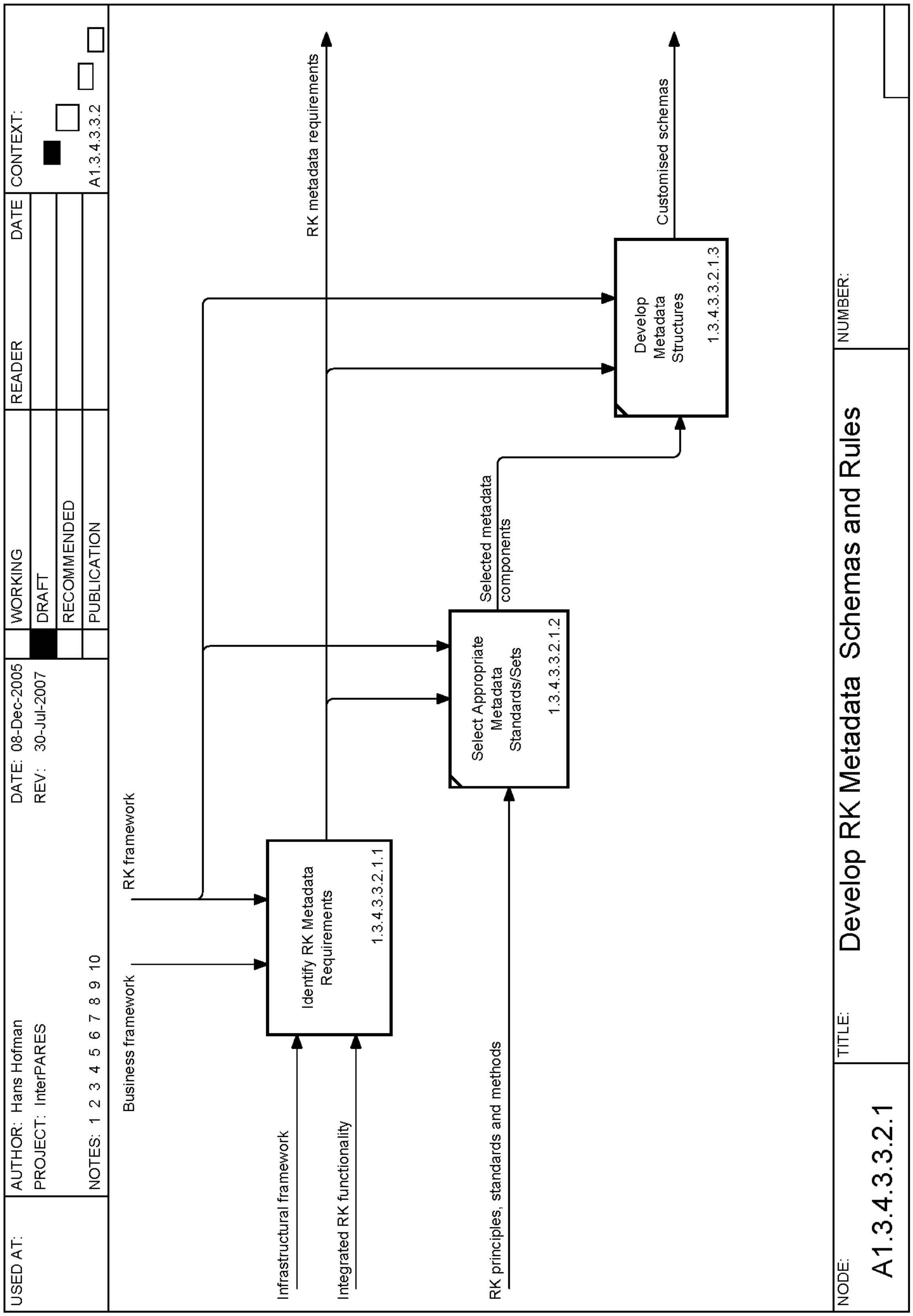

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.2.5
RK HRM instruments
Assigned RK responsibilities
Risk management framework
Preservation policy
Access framework
Evaluation of RK performance
Synthesised records requirements
Appraisal policy
Existing sets of functional RK requirements
Infrastructural framework
Analyse Different Sets of Requirements
1.3.4.3.2.5.3.1
Identified set of RK functional requirements
Compare identified RK Functional Requirements with Existing Standards/Sets
1.3.4.3.2.5.3.2
Existing standards or sets of functional requirements to be used
Define Set of RK Functional Requirements
1.3.4.3.2.5.3.3
Integrated RK functionality
NODE:
A1.3.4.3.2.5.3
TITLE:
Integrate Recordkeeping Functionality
NUMBER:

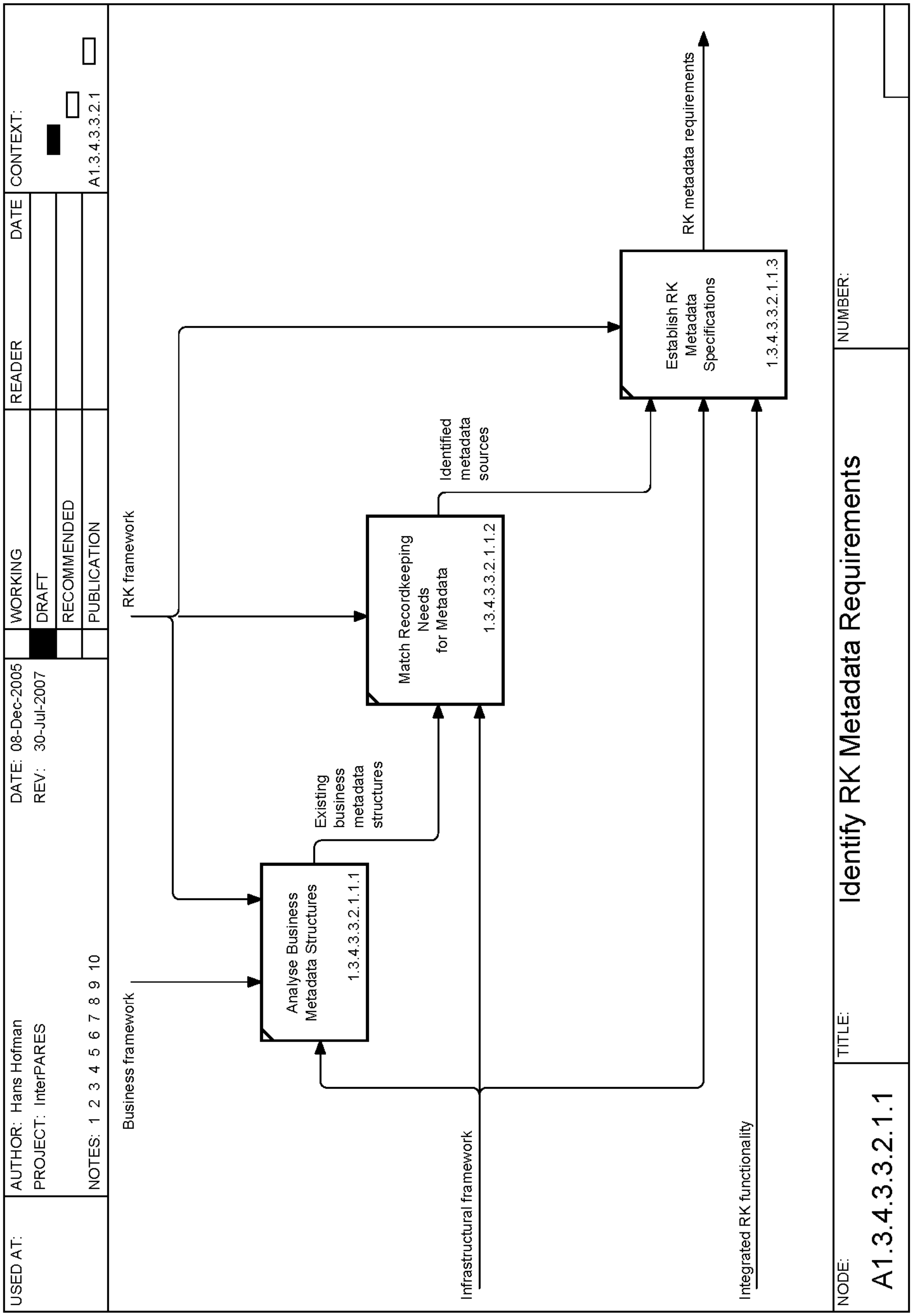

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
DATE: 08-Dec-2005
REV: 30-Jul-2007
NOTES: 1 2 3 4 5 6 7 8 9 10
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3
Financial plans
RK framework
Business framework
Functional analysis
State of technology
Evaluation of RK performance
RK principles, standards and methods
Integrated RK functionality
Retrieval requirements
Infrastructural framework
Manage RK Implementation
1.3.4.3.3.1
RK implementation performance criteria
RK implementation plans
Develop Recordkeeping Instruments
1.3.4.3.3.2
RK metadata requirements
Education program and courses
RK instruments
Design Recordkeeping Infrastructure
1.3.4.3.3.3
Train Staff
1.3.4.3.3.4
Monitor/Evaluate RK Implementation
1.3.4.3.3.5
RK implementation plans & progress reports
Requirements for RK in system(s)
Design of RK processes
Trained staff
RK implementation progress reports
NODE:
A1.3.4.3.3
TITLE:
Implement Recordkeeping Framework
NUMBER:

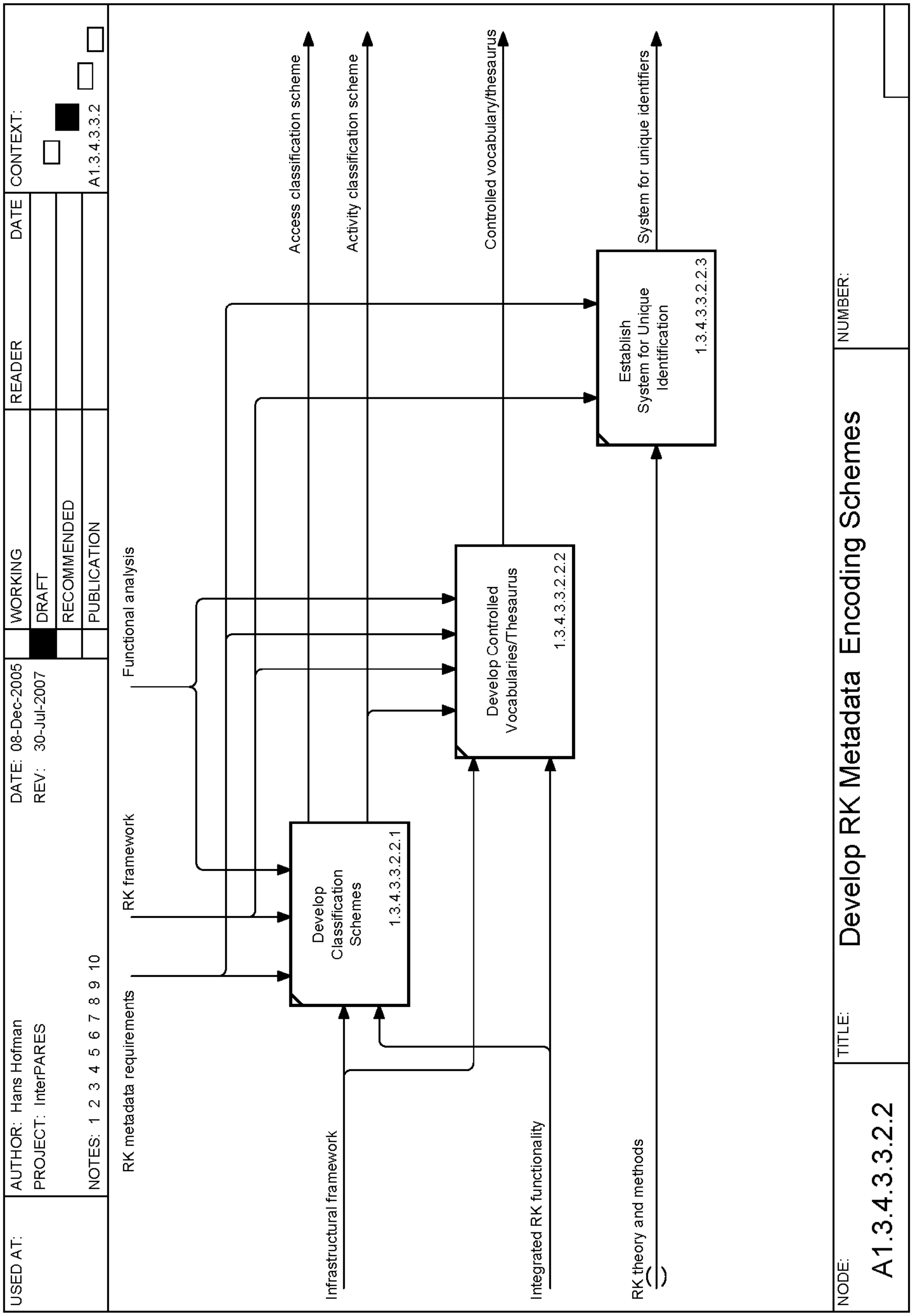

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.3
RK framework
Business framework
Financial plans
Capacity assessment
Evaluation of RK performance
Infrastructural framework
RK implementation progress reports
Analyse Readiness of Organisation/ Capacity Level
1.3.4.3.3.1.1
Design Transition Strategies
1.3.4.3.3.1.2
Transition strategies
Establish RK Implementation Plan
1.3.4.3.3.1.3
RK implementation plans
RK implementation performance criteria
NODE:
A1.3.4.3.3.1
TITLE:
Manage RK Implementation
NUMBER:

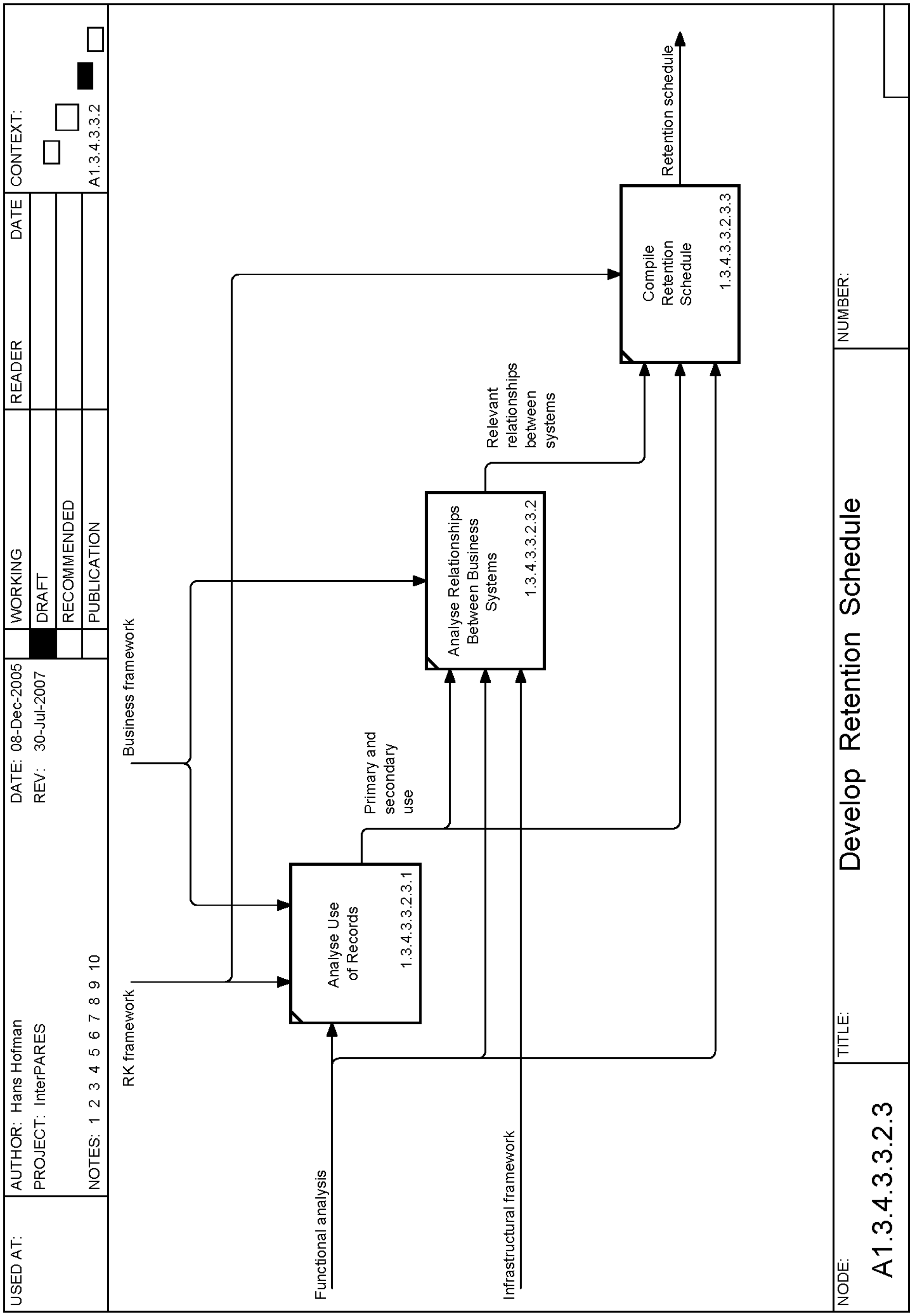

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.3.1
Transition strategies
Capacity assessment
Define RK Scope, Goals, Resources and Deliverables
1.3.4.3.3.1.3.1
Defined RK objectives and deliverables
RK framework
Financial plans
Define RK Implementation Plan
1.3.4.3.3.1.3.2
RK implementation progress reports
Evaluation of RK performance
RK implementation plans
Develop RK Implementation Performance Criteria
1.3.4.3.3.1.3.3
RK implementation performance criteria
NODE:
A1.3.4.3.3.1.3
TITLE:
Establish RK Implementation Plan
NUMBER:

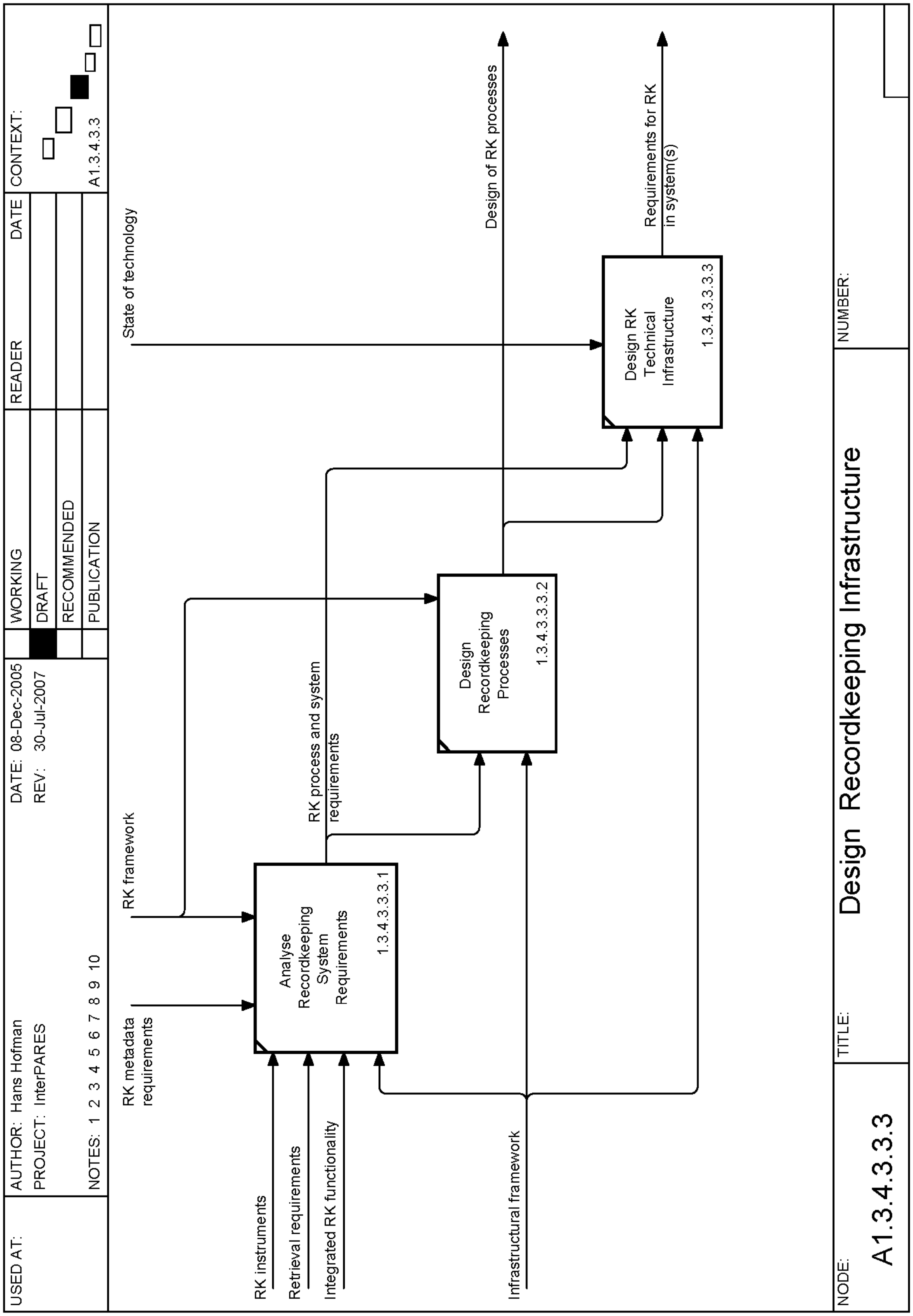

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.3
Business framework
RK principles, standards and methods
Integrated RK functionality
Infrastructural framework
Functional analysis
RK framework
Develop RK Metadata Schemas and Rules
1.3.4.3.3.2.1
Customised schemas
Develop RK Metadata Encoding Schemes
1.3.4.3.3.2.2
Access classification scheme
Controlled vocabulary/thesaurus
System for unique identifiers
Activity classification scheme
Develop Retention Schedule
1.3.4.3.3.2.3
Retention schedule
Develop RK Education Program and Courses
1.3.4.3.3.2.4
RK implementation plans
RK metadata requirements
RK instruments
Education program and courses
NODE:
A1.3.4.3.3.2
TITLE:
Develop Recordkeeping Instruments
NUMBER:

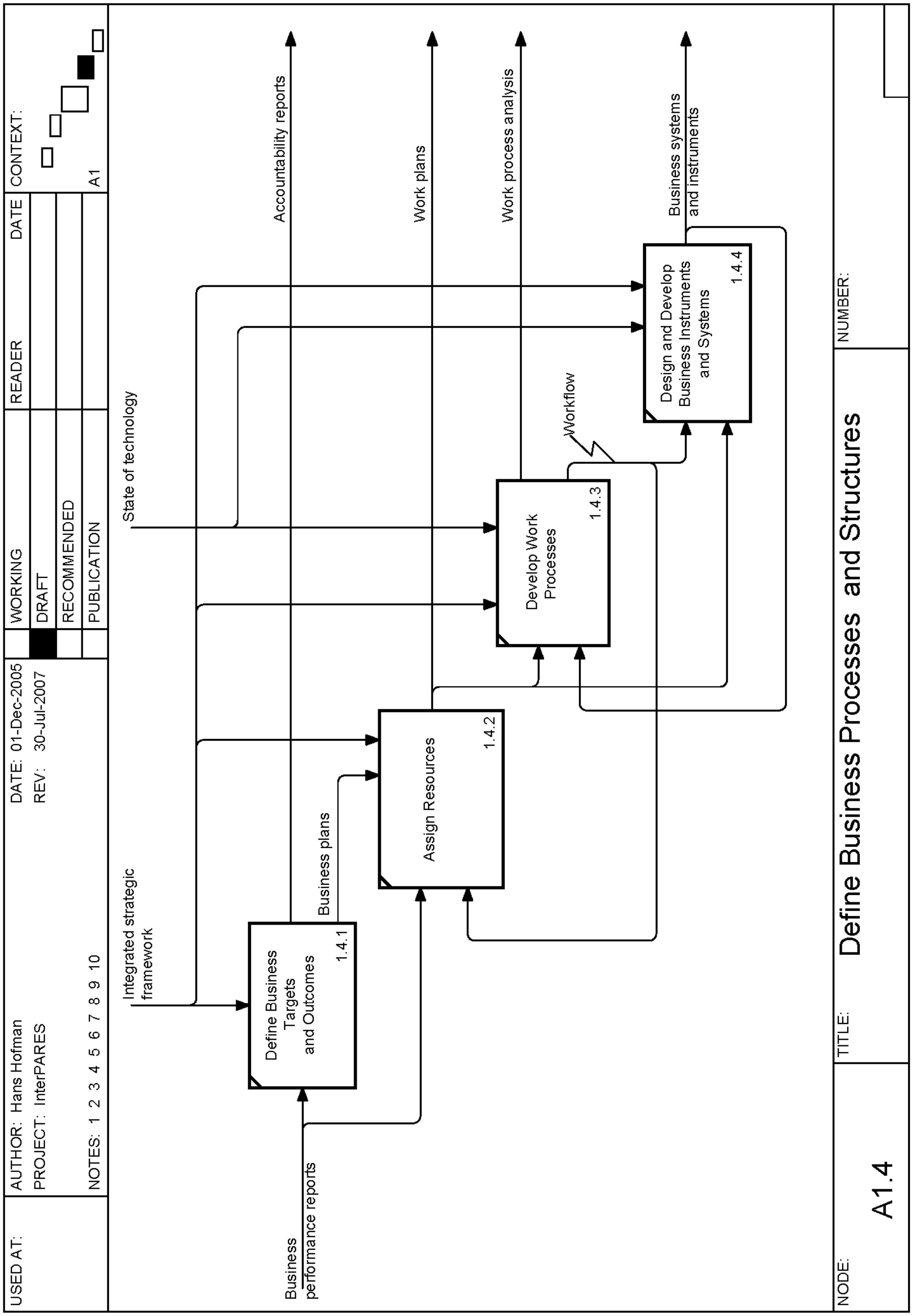

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.3.2
Infrastructural framework
Integrated RK functionality
Business framework
RK framework
Identify RK Metadata Requirements
1.3.4.3.3.2.1.1
RK principles, standards and methods
Select Appropriate Metadata Standards/Sets
1.3.4.3.3.2.1.2
Selected metadata components
RK metadata requirements
Develop Metadata Structures
1.3.4.3.3.2.1.3
Customised schemas
NODE:
A1.3.4.3.3.2.1
TITLE:
Develop RK Metadata Schemas and Rules
NUMBER:

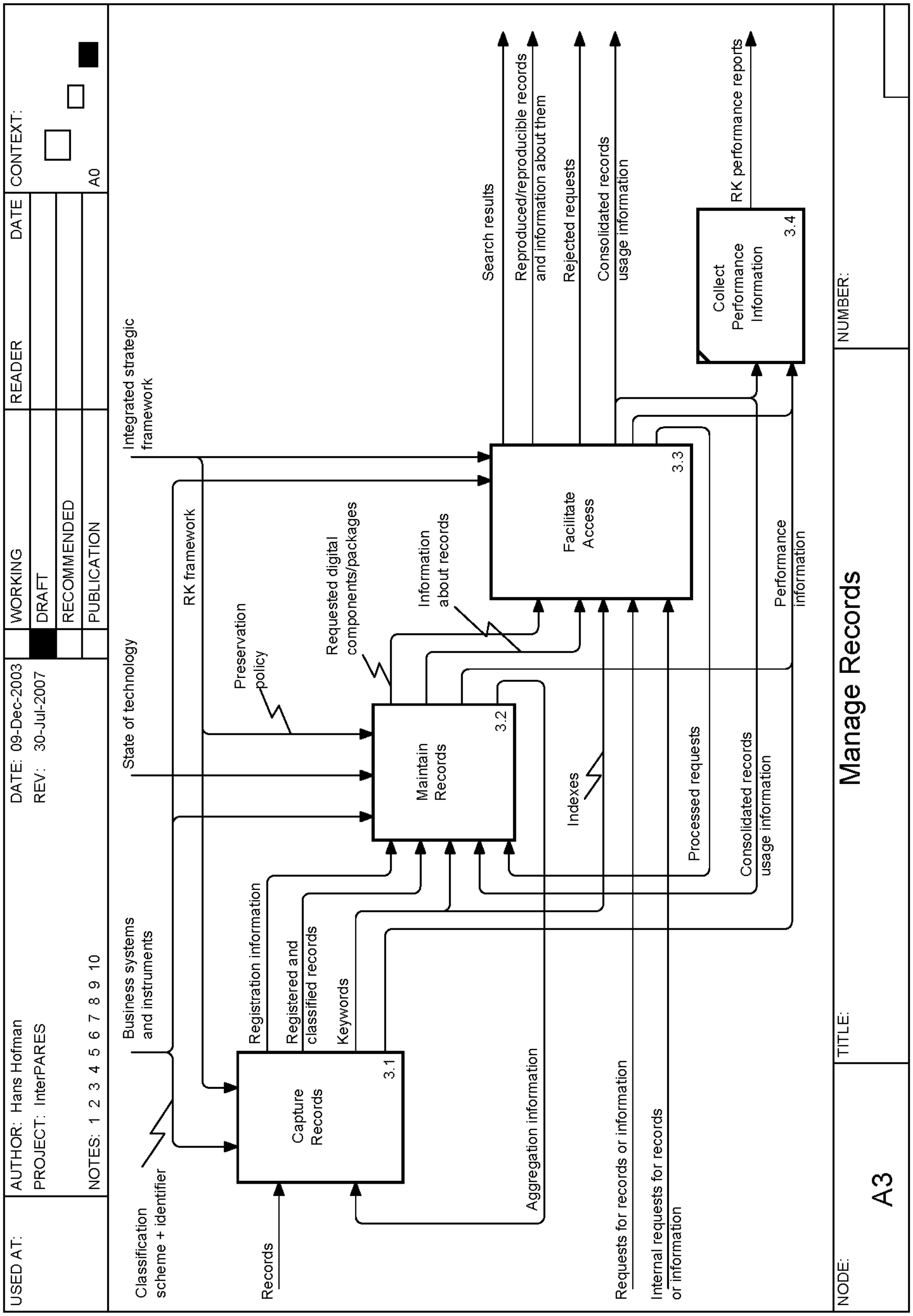

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.3.2.1
Business framework
RK framework
Infrastructural framework
Integrated RK functionality
Analyse Business Metadata Structures
1.3.4.3.3.2.1.1.1
Existing business metadata structures
Match Recordkeeping Needs for Metadata
1.3.4.3.3.2.1.1.2
Identified metadata sources
Establish RK Metadata Specifications
1.3.4.3.3.2.1.1.3
RK metadata requirements
NODE:
A1.3.4.3.3.2.1.1
TITLE:
Identify RK Metadata Requirements
NUMBER:

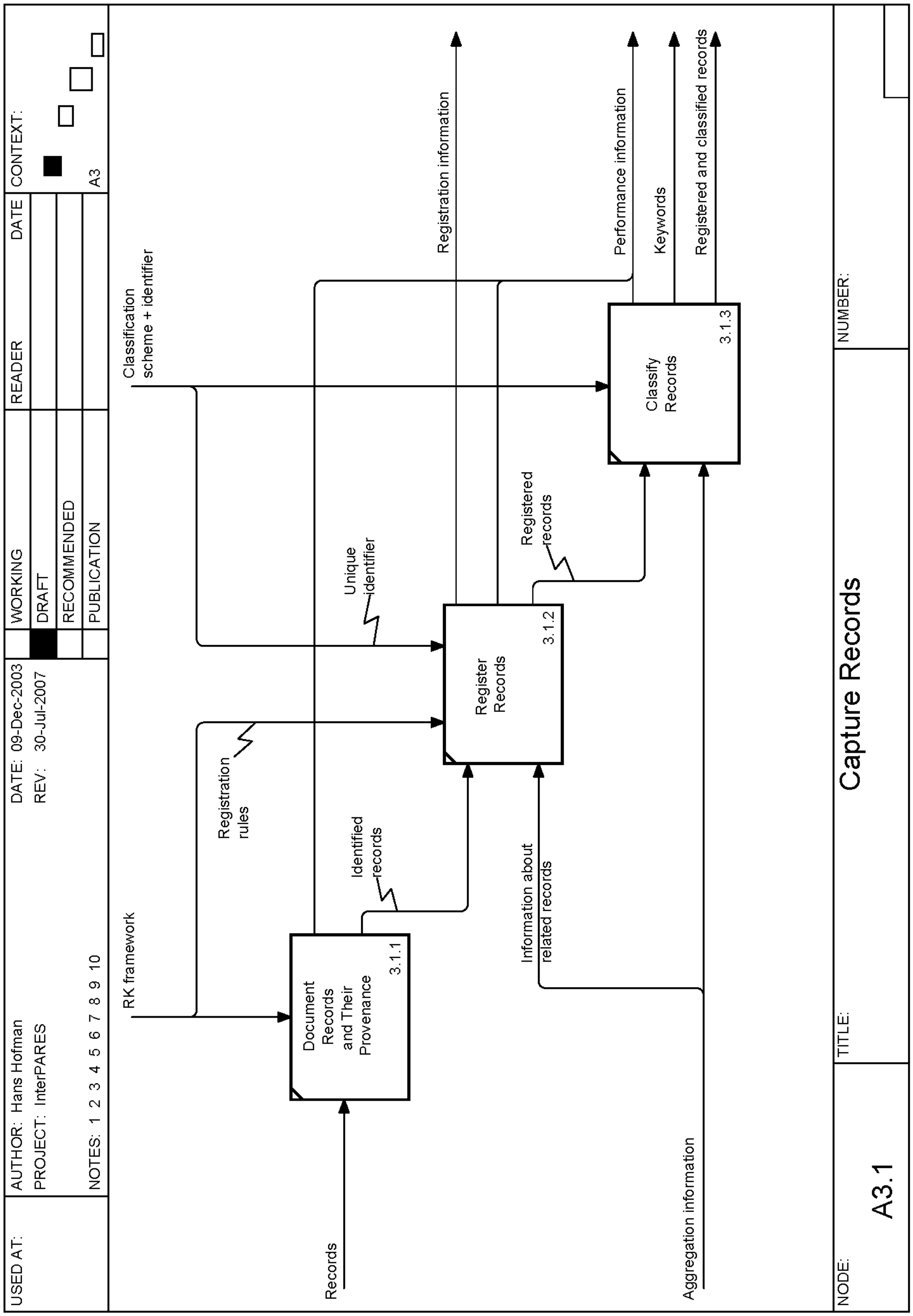

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.3.2
RK metadata requirements
RK framework
Functional analysis
Infrastructural framework
Integrated RK functionality
RK theory and methods
Develop Classification Schemes
1.3.4.3.3.2.2.1
Develop Controlled Vocabularies/Thesaurus
1.3.4.3.3.2.2.2
Establish System for Unique Identification
1.3.4.3.3.2.2.3
Access classification scheme
Activity classification scheme
Controlled vocabulary/thesaurus
System for unique identifiers
NODE:
A1.3.4.3.3.2.2
TITLE:
Develop RK Metadata Encoding Schemes
NUMBER:

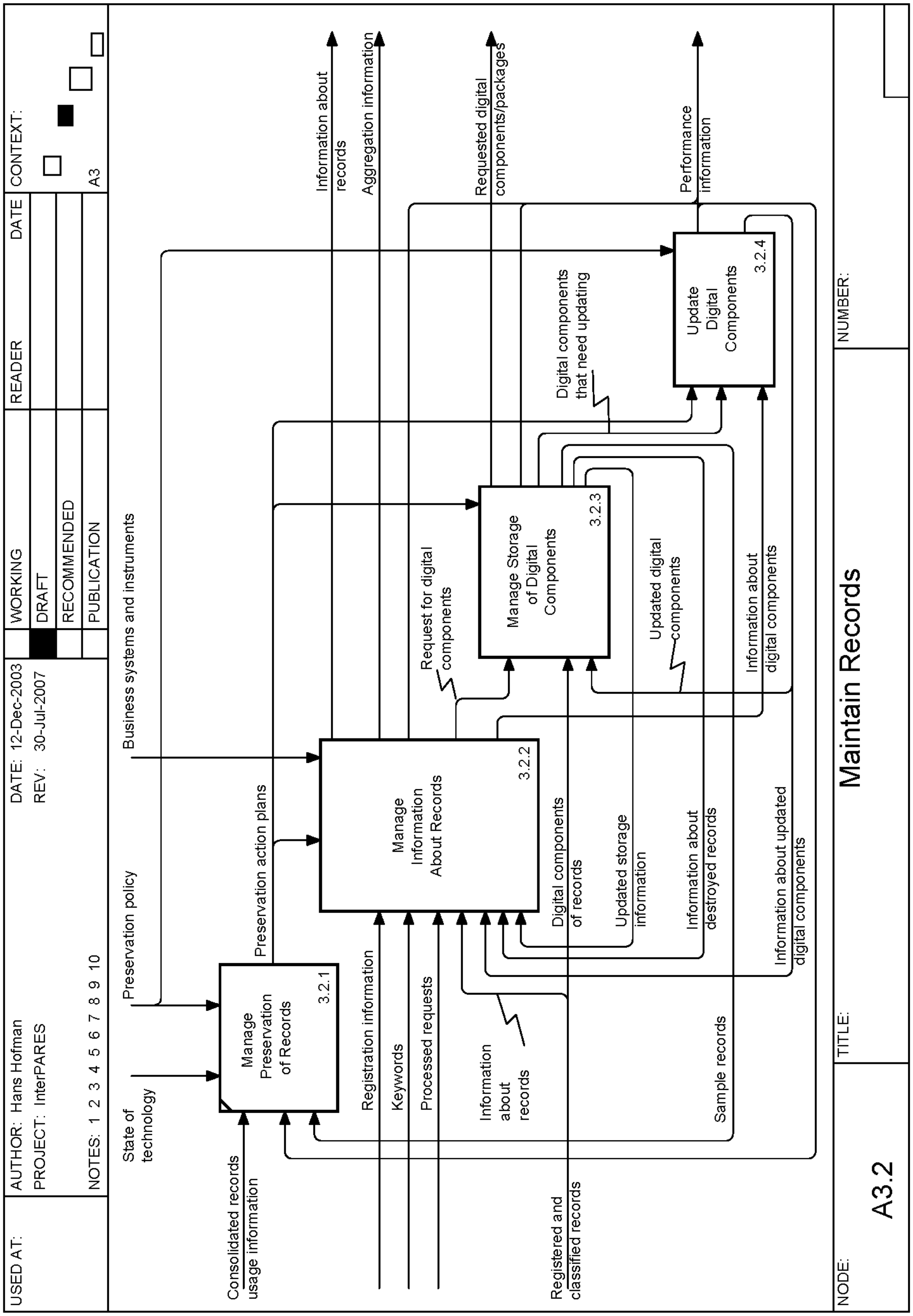

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.3.2
Functional analysis
RK framework
Business framework
Infrastructural framework
Analyse Use of Records
1.3.4.3.3.2.3.1
Primary and secondary use
Analyse Relationships Between Business Systems
1.3.4.3.3.2.3.2
Relevant relationships between systems
Compile Retention Schedule
1.3.4.3.3.2.3.3
Retention schedule
NODE:
A1.3.4.3.3.2.3
TITLE:
Develop Retention Schedule
NUMBER:

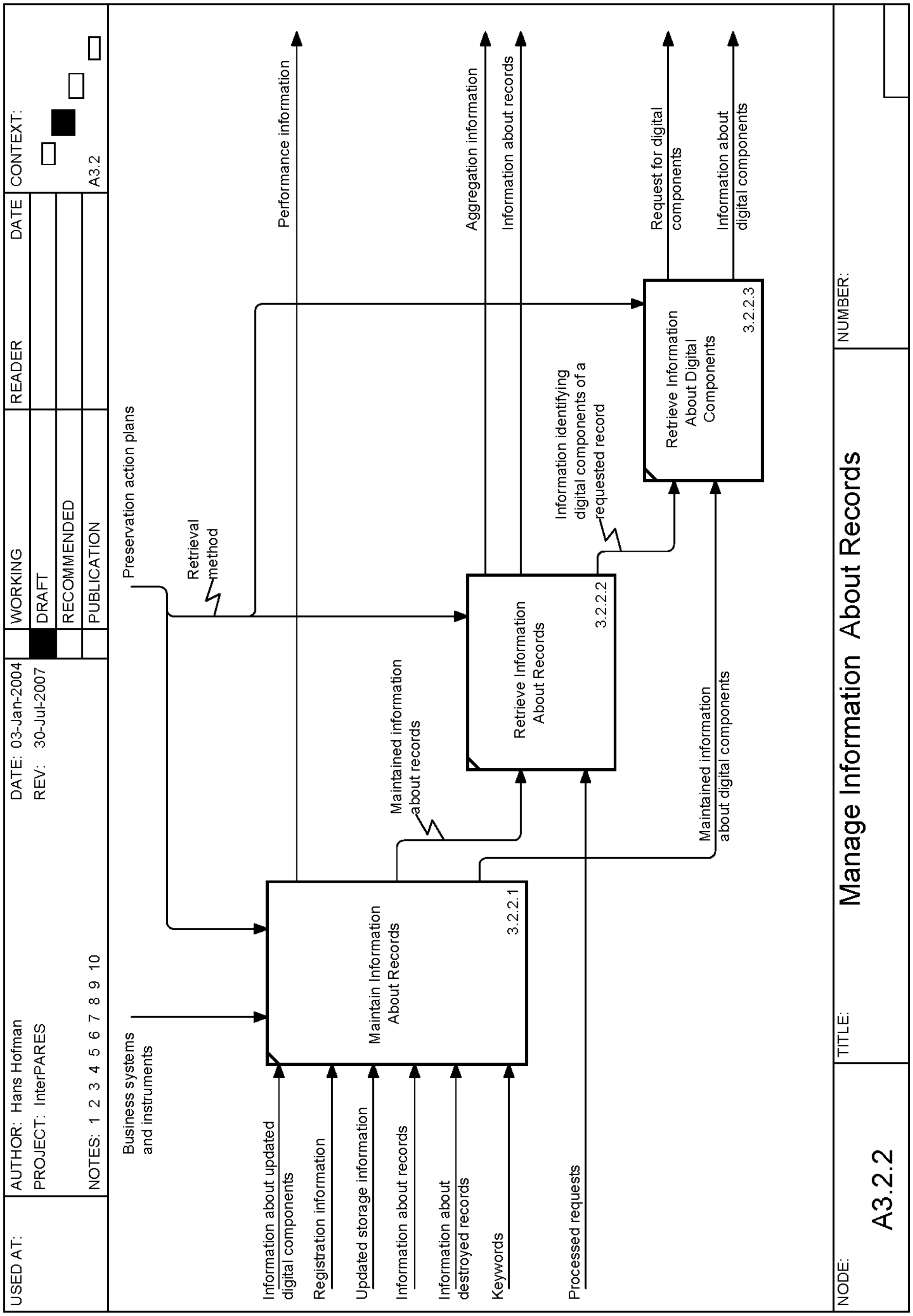

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 08-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1.3.4.3.3
RK metadata requirements
RK framework
RK instruments
Retrieval requirements
Integrated RK functionality
Infrastructural framework
Analyse Recordkeeping System Requirements
1.3.4.3.3.3.1
RK process and system requirements
Design Recordkeeping Processes
1.3.4.3.3.3.2
State of technology
Design of RK processes
Design RK Technical Infrastructure
1.3.4.3.3.3.3
Requirements for RK in system(s)
NODE:
A1.3.4.3.3.3
TITLE:
Design Recordkeeping Infrastructure
NUMBER:

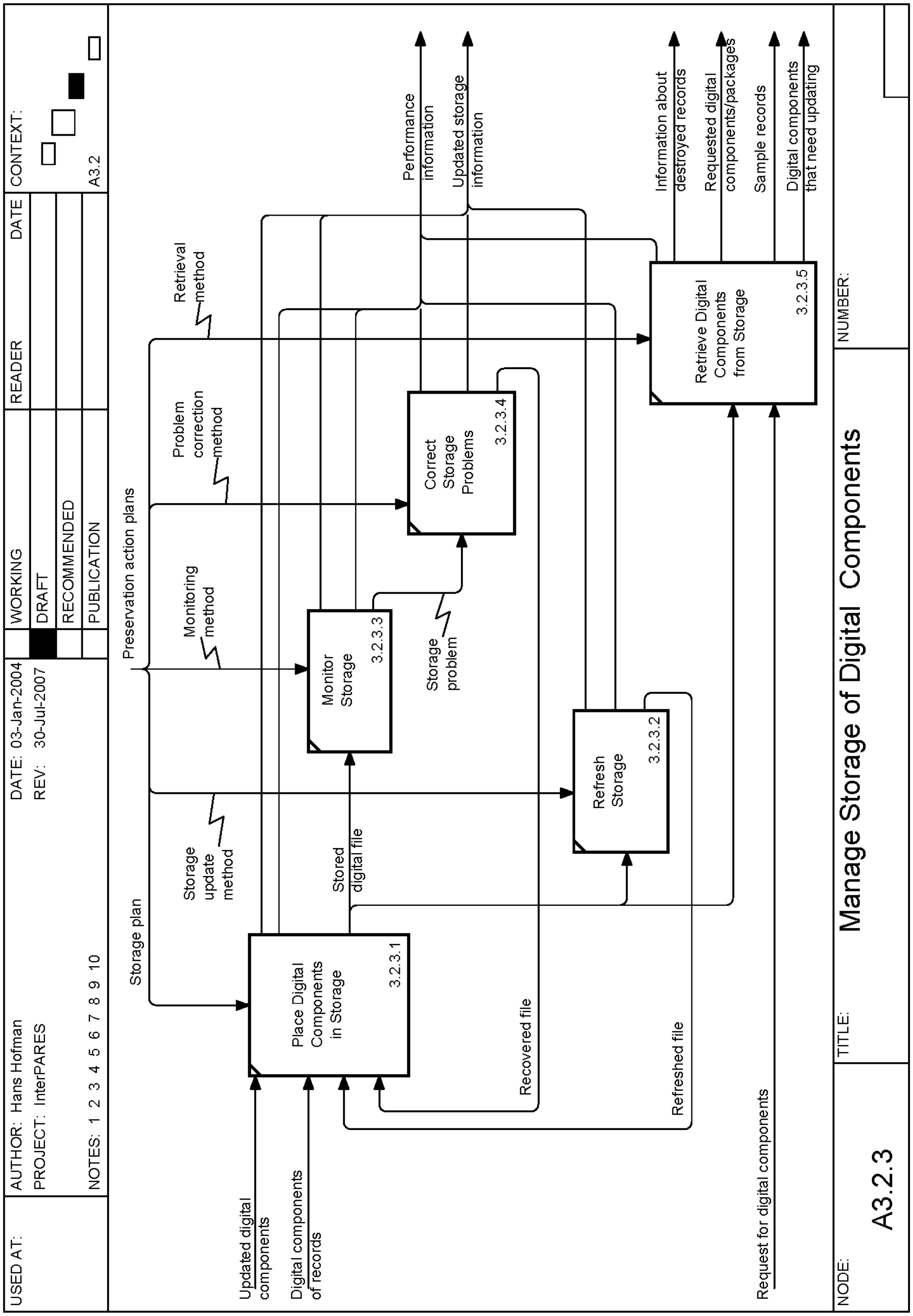

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 01-Dec-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A1
Integrated strategic framework
State of technology
Business performance reports
Define Business Targets and Outcomes
1.4.1
Business plans
Assign Resources
1.4.2
Develop Work Processes
1.4.3
Workflow
Design and Develop Business Instruments and Systems
1.4.4
Accountability reports
Work plans
Work process analysis
Business systems and instruments
NODE:
A1.4
TITLE:
Define Business Processes and Structures
NUMBER:

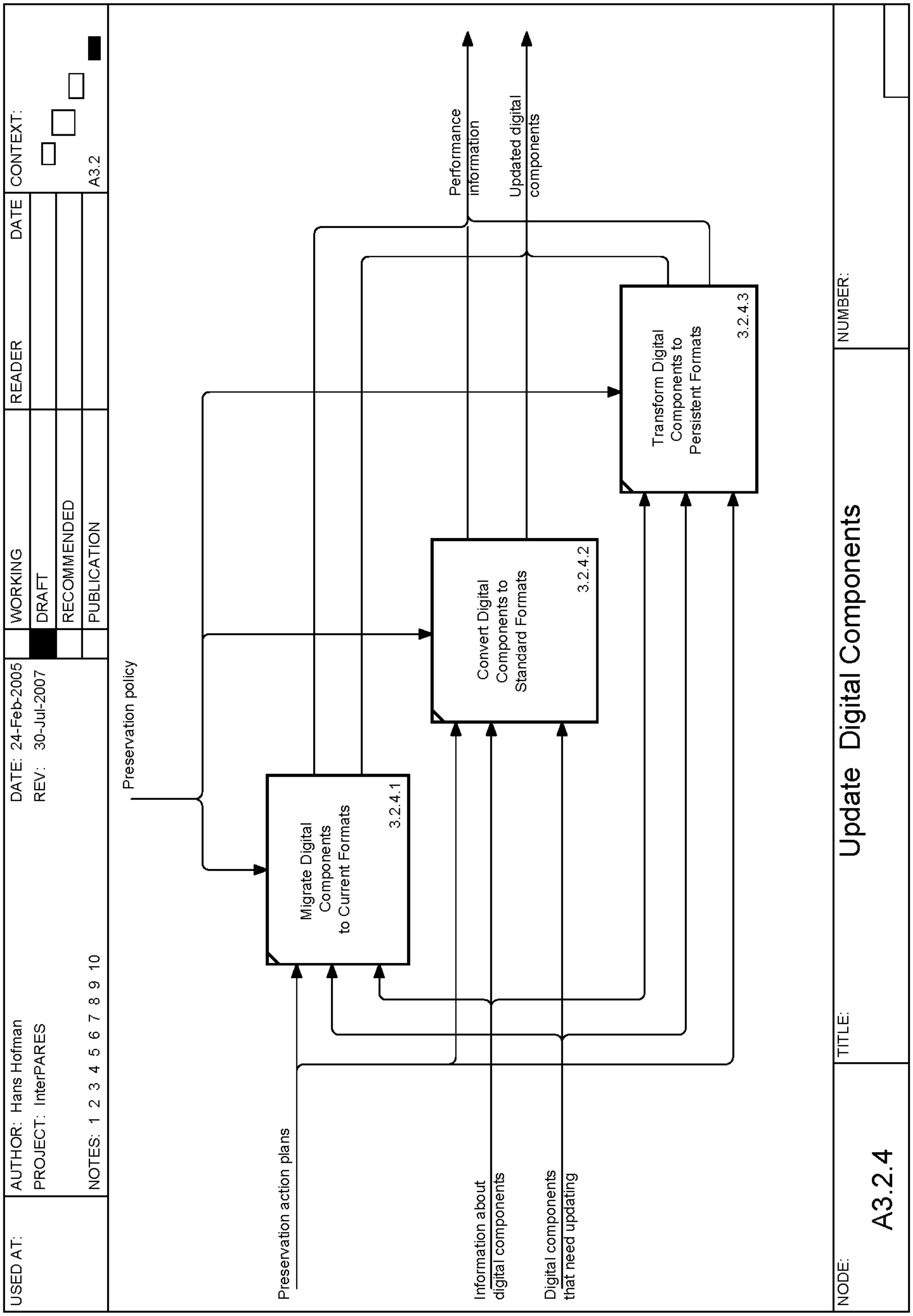

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 09-Dec-2003
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A0
Classification scheme + identifier
Records
Business systems and instruments
Capture Records
3.1
Aggregation information
Registration information
Registered and classified records
Keywords
State of technology
RK framework
Preservation policy
Maintain Records
3.2
Requested digital components/packages
Information about records
Indexes
Processed requests
Consolidated records usage information
Integrated strategic framework
Facilitate Access
3.3
Performance information
Search results
Reproduced/reproducible records and information about them
Rejected requests
Consolidated records usage information
Collect Performance Information
3.4
RK performance reports
Requests for records or information
Internal requests for records or information
NODE:
A3
TITLE:
Manage Records
NUMBER:

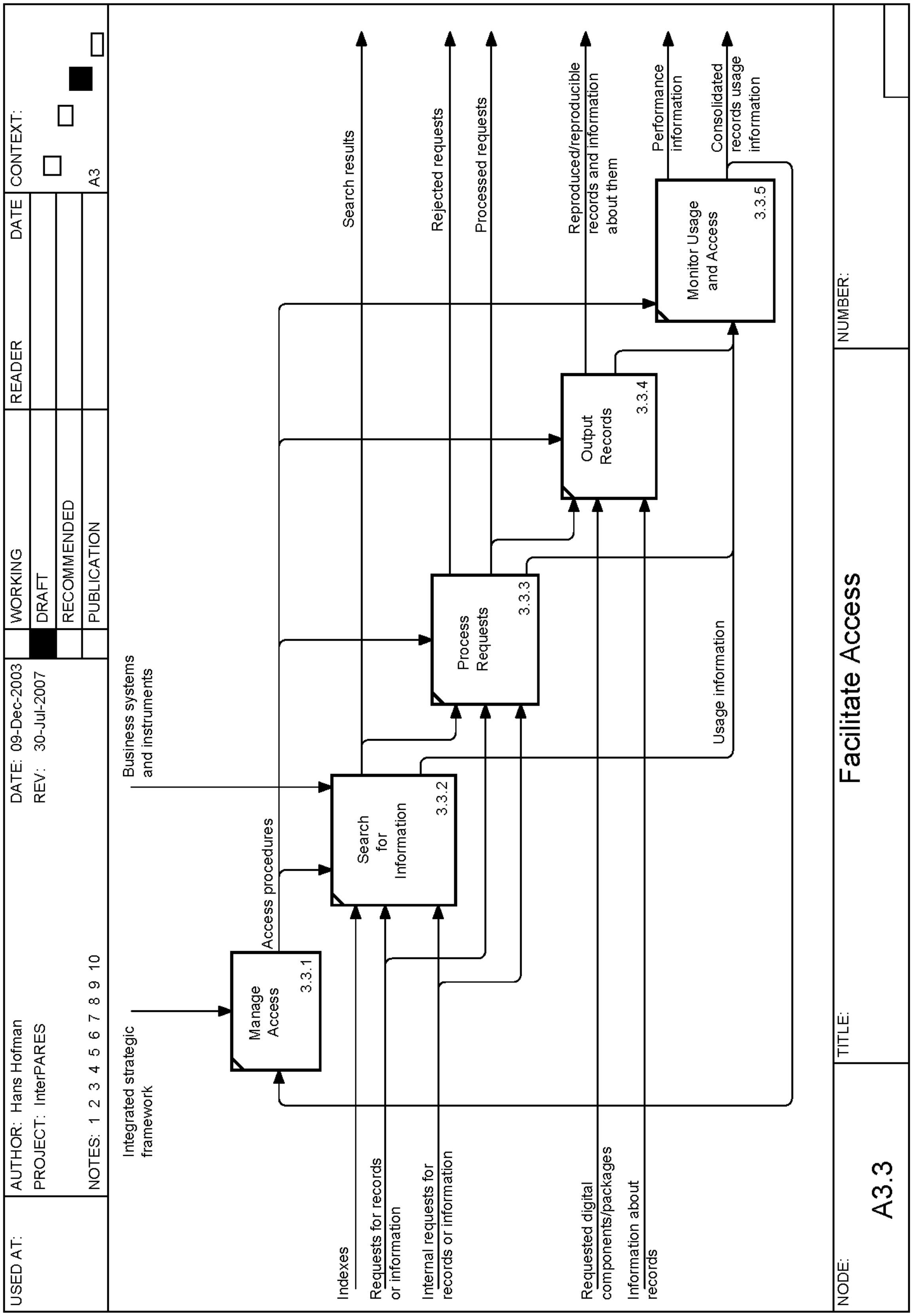

USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 09-Dec-2003
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3
Records
RK framework
Document Records and Their Provenance
3.1.1
Identified records
Registration rules
Unique identifier
Register Records
3.1.2
Information about related records
Registration information
Classification scheme + identifier
Registered records
Classify Records
3.1.3
Performance information
Keywords
Registered and classified records
Aggregation information
NODE:
A3.1
TITLE:
Capture Records
NUMBER:


USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 12-Dec-2003
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3
Manage Preservation of Records
3.2.1
Manage Information About Records
3.2.2
Manage Storage of Digital Components
3.2.3
Update Digital Components
3.2.4
State of technology
Preservation policy
Business systems and instruments
Preservation action plans
Consolidated records usage information
Registered and classified records
Registration information
Keywords
Processed requests
Information about records
Sample records
Digital components of records
Updated storage information
Information about destroyed records
Information about updated digital components
Request for digital components
Updated digital components
Information about digital components
Digital components that need updating
Information about records
Aggregation information
Requested digital components/packages
Performance information
NODE:
A3.2
TITLE:
Maintain Records
NUMBER:


USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 03-Jan-2004
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3.2
Business systems and instruments
Preservation action plans
Retrieval method
Performance information
Aggregation information
Information about records
Request for digital components
Information about digital components
Maintain Information About Records
3.2.2.1
Retrieve Information About Records
3.2.2.2
Retrieve Information About Digital Components
3.2.2.3
Maintained information about records
Maintained information about digital components
Information identifying digital components of a requested record
Information about updated digital components
Registration information
Updated storage information
Information about records
Information about destroyed records
Keywords
Processed requests
NODE:
A3.2.2
TITLE:
Manage Information About Records
NUMBER:


USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 03-Jan-2004
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3.2
Place Digital Components in Storage
3.2.3.1
Refresh Storage
3.2.3.2
Monitor Storage
3.2.3.3
Correct Storage Problems
3.2.3.4
Retrieve Digital Components from Storage
3.2.3.5
Storage plan
Preservation action plans
Storage update method
Monitoring method
Problem correction method
Retrieval method
Stored digital file
Storage problem
Recovered file
Refreshed file
Updated digital components
Digital components of records
Request for digital components
Performance information
Updated storage information
Information about destroyed records
Requested digital components/packages
Sample records
Digital components that need updating
NODE:
A3.2.3
TITLE:
Manage Storage of Digital Components
NUMBER:


USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 24-Feb-2005
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3.2
Preservation policy
Preservation action plans
Information about digital components
Digital components that need updating
Migrate Digital Components to Current Formats
3.2.4.1
Convert Digital Components to Standard Formats
3.2.4.2
Transform Digital Components to Persistent Formats
3.2.4.3
Performance information
Updated digital components
NODE:
A3.2.4
TITLE:
Update Digital Components
NUMBER:


USED AT:
AUTHOR: Hans Hofman
PROJECT: InterPARES
NOTES: 1 2 3 4 5 6 7 8 9 10
DATE: 09-Dec-2003
REV: 30-Jul-2007
WORKING
DRAFT
RECOMMENDED
PUBLICATION
READER
DATE
CONTEXT:
A3
Integrated strategic framework
Business systems and instruments
Access procedures
Indexes
Requests for records or information
Internal requests for records or information
Requested digital components/packages
Information about records
Manage Access
3.3.1
Search for Information
3.3.2
Process Requests
3.3.3
Output Records
3.3.4
Monitor Usage and Access
3.3.5
Usage information
Search results
Rejected requests
Processed requests
Reproduced/reproducible records and information about them
Performance information
Consolidated records usage information
NODE:
A3.3
TITLE:
Facilitate Access
NUMBER:

Business-driven Recordkeeping Model Activity Definitions

A0, Manage Business

Under the control of organisational/business, juridical and legal requirements, and limited by the actual capacity of the organisation and the possibilities available within the State of the Art of Information Technology, create and manage records as long as required in order to enable and support one or more identified business activities as well as to meet the applicable juridical requirements.

A1, Manage Business Framework

Establish a framework for records creation and management in an organisation, in line with its business needs, implement it based on the capability level that is identified for the organisation, monitor and evaluate the performance and application of the established framework both for the business activities and recordkeeping processes, and if necessary subsequently adjust the framework to the needs and level required.

A1.1, Analyse Needs and Risks

Establish the business needs for records and identify the risks if these records will not be properly created and managed in relation to the business context to identify the requirements for recordkeeping. This activity is controlled by legal, juridical and organisational or business requirements.

A1.2, Establish Governance

Set the overall strategic direction of an organisation by establishing the set and levels of responsibilities and practices necessary to ensure that the organisation is accountable for fulfilling its mandate, complying to its legal obligations, achieving its stakeholders' objectives, meeting the current societal, ethical and moral duties, managing its risks appropriately, using its resources responsibly and monitoring its performance effectively.

A1.3, Manage Strategic Framework

Establish strategic plans outlining the organisation's current and future direction, priorities and resource allocation strategies, in line with its business needs and key stakeholder interests, as well as including the required mitigation of business risks identified, implement them within an overall strategic framework, monitor the performance and application of the established plans both for the business activities and recordkeeping processes and, if necessary, subsequently adjust the plans to continue to meet business and key stakeholder needs and interests.

A1.3.1, Analyse Mandates

Critically evaluate an organisation's external mandate, identify the responsibilities involved within the context of the organisation's corporate culture and accountability framework, and subsequently identify the key mandate drivers that will guide, support and control the organisation's activities.

A1.3.2, Analyse Stakeholder Interests

Collect, evaluate input from key stakeholders (individuals, groups, other organisations, etc.) and identify their interests with respect to the organisation's business activities and objectives, its current and future direction, and its operational priorities and outcomes.

A1.3.3, Analyse Market

Collect and evaluate economic, operational, and/or competitive performance data regarding the business sector(s) within which the organisation operates for the purpose of identifying the products and/or services the organisation should produce and provide, the needs and interests of its customers, and the strengths and weaknesses of its competition. A market analysis provides an organisation with the baseline information it needs to develop and manage its operations to best achieve its goals and enhance its competitive success.

A1.3.4, Develop Strategic Framework

Develop, implement and manage the different (strategic) specific frameworks and integrate them into one strategic framework, based on the outcome of the analysis of the mandate(s), the stakeholders' interests and the market. The strategic framework will provide the guidance and control needed to coordinate the key risk management, business and recordkeeping functions of the organisation.

A1.3.4.1, Develop Risk Management Framework

Develop, implement and manage a comprehensive administrative and operational framework for managing identified or potential risks based upon the analysis of the juridical requirements, the organization's responsibilities, societal influences, stakeholder interests, mandate drivers and market analyses within the constraints of the organisation's business.

A1.3.4.2, Develop Business Framework

Develop, implement and manage a comprehensive administrative and operational framework that will guide and control an organisation's business activities, within the constraints of the juridical requirements, the existing corporate culture and the risk management framework, through consideration of recordkeeping implementation plans and progress reports, the business infrastructural framework, mandate analysis, market analysis, stakeholder interest information, and business performance analyses. The business framework will be the basis for accountability reports and be accompanied by financial plans.

A1.3.4.2.1, Analyse External and Internal Drivers and Constraints

Collect and evaluate information about the key internal and external operational, administrative, cultural, documentary and resource allocation factors that may impact upon an organisation's ability to carry out its business, to meet its purpose and to achieve its outcomes, in order to provide a contextual analysis for establishing an organisation's business framework.

A1.3.4.2.2, Define Business Framework

Design and specify a comprehensive operational and administrative structure for guiding and overseeing an organisation's business activities, within the constraints of the existing corporate culture and the risk management framework, and based upon the thorough consideration of the contextual analysis, the recordkeeping implementation plans, the recordkeeping infrastructure framework, and performance analyses.

A1.3.4.2.3, Monitor and Evaluate Business Framework

Periodically assess whether the way an organisation's business framework is operationalised and still appropriate based upon analysis of business performance information in relation to the existing risk management framework and the contextual analysis of key drivers and constraints. Based on the monitoring, produce reports to inform the define business framework function to confirm or revise the business framework.

A1.3.4.3, Develop RK Framework

Develop, implement and manage a comprehensive administrative and operational structure that will guide and control an organisation's recordkeeping activities, within the constraints of the current state of technology, the organisation's existing corporate culture, its risk management and business frameworks, the financial plan, and its juridical requirements, and taking into account performance reports on business and recordkeeping activities.

A1.3.4.3.1, Analyse Recordkeeping Needs and Risks

Identify, describe, document and analyse an organisation's work processes ('sequential analysis'), identify what records are needed to adequately document and support those processes, and identify potential associated risks in relation to performing the recordkeeping function.

A1.3.4.3.1.1, Identify/Describe Business Processes

Given the relevant juridical, legal and organisational requirements, and past information about business and recordkeeping performance, identify the requirements for retention of the records created.

A1.3.4.3.1.2, Identify Business Need for Records

Identify what records are needed for the business processes at transaction level based on a risk analysis of these processes in connection to legal, business, organisational and societal requirements.

A1.3.4.3.1.2.1, Analyse Legal and Juridical Requirements in Relation to Business Activities

For the business processes/activities, determine the related legal and juridical requirements for records creation and the risks of not meeting these requirements.

A1.3.4.3.1.2.2, Identify Records to be Created

Given the relevant juridical, legal and organisational requirements, and past information about business and recordkeeping performance, identify what records should be created, their structure and form, and with what technologies.

A1.3.4.3.1.2.3, Identify the Retention Requirements for the Records to be Created

Given the relevant juridical, legal and organisational requirements, and past information about business and recordkeeping performance, identify the requirements for retention of the records created.

A1.3.4.3.1.2.4, Derive Requirements for Authenticity

Given the requirements for the records to be created in the different business processes specify, if possible, the characteristics of those records that are essential to their intent and the message they are supposed to convey in the given business context.

A1.3.4.3.1.3, Identify Risks from RK Perspective

Based on a functional analysis of an organisation's business processes, as well as consideration of the organisation's mandate drivers, the business need for records and the requirements for retention, identify the potential associated risks in relation to performing the recordkeeping function.

A1.3.4.3.2, Define Recordkeeping Framework

Design and specify a comprehensive operational and administrative structure for guiding and overseeing the recordkeeping function in the organisation a) within the constraints of the business framework, the financial plans, the current state of technology, the organisation's existing corporate culture, and b) based upon an analysis of the business needs for records and the risk analysis, identification of the requirements for creating and managing records and the related functionality in systems, assign who will be responsible for the different recordkeeping roles, establish policies for retention, appraisal, preservation and access and determine criteria for evaluation of the performance.

A1.3.4.3.2.1, Synthesise Requirements for Records

Synthesise the requirements that will ensure the authenticity, reliability, usability and integrity of records, within the given business context, including the requirements for organizing and presenting records and/or their aggregates.

A1.3.4.3.2.1.1, Derive Record-type Properties Critical for Authenticity

Determine the structure and documentary form of the different types of records required to carry out the various business functions that are essential to their authenticity in the given business context, including their technical characteristics.

A1.3.4.3.2.1.2, Determine Guidelines for Organising Records and for Metadata Framework

Determine the rules and guidelines for classifying records, identify what metadata standards to follow and the requirements for developing metadata schema in relation to the business and recordkeeping activities, and the levels of aggregation needed.

A1.3.4.3.2.1.3, Determine Requirements for Presenting Records/Aggregates

Analyse, define and document the requirements for presenting records and their aggregates, based upon the guidelines for organizing them, so they can be presented to fulfill requests for records or record aggregates in ways that reflects their interrelationships.

A1.3.4.3.2.1.4, Synthesise Requirements for Records/Aggregates

Combine the set of recordkeeping framework requirements that determine the type of records, their characteristics critical for authenticity, the rules for organising and presenting records and their aggregates and rules and requirements for the required metadata.

A1.3.4.3.2.2, Define Appraisal Policy

Determine the method and rules for appraisal of records based upon the retention requirements identified in the risk analysis, the infrastructural framework, and upon the business, legal, organisational, and societal requirements.

A1.3.4.3.2.2.1, Manage Appraisal Policy

Provide overall control and co-ordination of an organisation's appraisal policy function via guidelines and directives issued in response to information received from the monitor/evaluate function.

A1.3.4.3.2.2.2, Develop Appraisal Policy

In response to the appraisal guidelines and directives, and based on consideration of the business framework, the organisation's synthesised records needs and risks requirements, its recordkeeping risk analysis, risk management plan, and its consolidated records usage information, define and document the guidelines, methods and rules that will make up the appraisal policy.

A1.3.4.3.2.2.3, Monitor/Evaluate Appraisal Policy

Periodically assess whether the way the appraisal policy is formulated and carried out in the organisation is still appropriate and in line with the business framework, and retention requirements based upon evaluation of recordkeeping performance in relation to the existing risk management assessment and appraisal policy directives. Based on the monitoring, produce reports to inform the appraisal policy management function to confirm or revise the appraisal policy.

A1.3.4.3.2.3, Define Preservation Policy

Determine the requirements, methods and rules for preserving records and related digital components within the framework of the existing business needs, the appraisal policy, the state of the art of technology, and the synthesised record requirements, and based upon the infrastructural framework and evaluation information of recordkeeping performance.

A1.3.4.3.2.3.1, Manage Preservation Policy

Based on critical characteristics of the created record types, the appraisal policy, the existing infrastructural framework, and the given state of the art of technology, determine the preservation requirements and preservation principles, rules and methods.

A1.3.4.3.2.3.2, Develop Preservation Policy

Develop the guidelines, methods and rules for preservation based on the identified preservation requirements, the appraisal policy, the essential characteristics of record types, and the evaluation of the preservation policy, and subsequent evaluate the most suitable preservation strategies, and the organisation's given infrastructural framework and the current state of the art of technology.

A1.3.4.3.2.3.2.1, Develop Experiments for Evaluating Most Suitable Preservation Strategies

Design experiments for evaluating the most suitable preservation strategies based on the essential characteristics of the different types of records (including their digital components), a sample of relevant record types, the preservation principles and the preservation requirements.

A1.3.4.3.2.3.2.2, Evaluate Potential Preservation Methods

Conduct experiments on samples of record types to assess the most suitable preservation strategy for each type of record (and their digital components) given the state of technology.

A1.3.4.3.2.3.2.3, Compile Preservation Policy

Synthesise information about targeted preservation methods with information from preservation policy evaluation reports, check this with the given appraisal policy, and reconcile these with identified preservation principles, the preservation requirements, and technological constraints to formulate a coordinated preservation policy.

A1.3.4.3.2.3.3, Monitor/ Evaluate Preservation Policy

Evaluate the suitability of the preservation policy based upon the information about the recordkeeping performance, the actual preservation policy and the needs of the business activities

A1.3.4.3.2.4, Define Access Framework

Determine the use and outreach within the given business context, identify the targeted communities and users, and their requirements for access and use of the records and develop and manage a comprehensive administrative and operational structure that will guide and control the access to records.

A1.3.4.3.2.4.1, Establish Retrieval Requirements

Based on prior usage, organisational requirements/business framework, and appraisal policy, specify retrieval requirements.

A1.3.4.3.2.4.2, Establish RK Security Policy

Determine objectives, methods, and rules for recordkeeping security including access for authorised individuals, denial of access for unauthorised individuals, data integrity, and auditability of access and violation of access.

A1.3.4.3.2.4.3, Establish Privacy Policy

Determine what privacy rules are valid and needed given the juridical requirements, the risk management framework, and information about the kinds of information in business records.

A1.3.4.3.2.4.4, Establish IPR Policy

Determine the rules for IPR and copy rights with respect to records given the business activities, the risk management framework, and legal and organisational requirements.

A1.3.4.3.2.4.5, Identify Access Rights

Given a security policy, IPR policy, privacy policy and recordkeeping security policy, identify privileges/rights for access to records and recordkeeping functions.

A1.3.4.3.2.4.6, Consolidate Access Framework

Consolidate retrieval requirements, privacy policy, Intellectual Property Rights (IPR) policy, and access privileges into a coherent framework and reconcile this with the identified reach within the given business context, and the targeted communities and users, and their requirements for access and use of the records.

A1.3.4.3.2.5, Specify Recordkeeping Framework

Establish 1) the recordkeeping framework's overall design, structure and integrated functionality, 2) the recordkeeping job functions, competencies and staff education/training instruments needed to implement, use and maintain the recordkeeping functions, and 3) the criteria against which to measure performance of the recordkeeping functions within the framework.

A1.3.4.3.2.5.1, Assign Recordkeeping Responsibilities

Identify the different roles with respect to recordkeeping and assign responsibilities to them, based upon the business and risk management frameworks and the established appraisal, preservation and access policies

A1.3.4.3.2.5.2, Develop Recordkeeping Competencies/Job Descriptions

Identify 1) competencies for recordkeeping functions, 2) develop job descriptions and 3) develop instruments that will support increasing the capabilities in staff with assigned recordkeeping responsibilities.

A1.3.4.3.2.5.3, Integrate Recordkeeping Functionality

Analyze various sets of records requirements and recordkeeping policies, translate them into functional requirements and design integrated recordkeeping functionality based upon the infrastructural framework, the risk management framework and the assigned responsibilities.

A1.3.4.3.2.5.3.1, Analyse Different Sets of Requirements

Analyse preservation policy, appraisal policy, synthesised record requirements and access framework to produce a requirements analysis document also taking into account recordkeeping HRM instruments, the recordkeeping framework and the risk management framework.

A1.3.4.3.2.5.3.2, Compare identified RK Functional Requirements with Existing Standards/Sets

Compare identified functional recordkeeping requirements with existing standards/sets of requirements, taking into account the infrastructural framework

A1.3.4.3.2.5.3.3, Define Set of RK Functional Requirements

Compile/define an integrated set of recordkeeping functional requirements based on the comparison with existing standards and the identified functional requirements needed to support the recordkeeping framework and processes.

A1.3.4.3.2.5.4, Compile Recordkeeping Framework

Compile the integrated recordkeeping functionality, preservation policy, appraisal policy and evaluation of recordkeeping performance to produce a comprehensive recordkeeping framework with assigned responsibilities and associated recordkeeping competencies

A1.3.4.3.3, Implement Recordkeeping Framework

Manage the implementation of the recordkeeping framework in line with the capacity of the organisation, carry out the actual implementation and monitor and evaluate its progress. This is a continuous process that will be influenced by the changes in business function(s) or by changes in legal juridical or technological circumstances.

A1.3.4.3.3.1, Manage RK Implementation

Analyze the readiness of the organisation and its capacity level, design the transition strategies and establish implementations plans and monitor their progress.

A1.3.4.3.3.1.1, Analyse Readiness of Organisation/ Capacity Level

Analyse the business and recordkeeping frameworks in close relation with the recordkeeping performance, information about business performance and implementation progress reports in order to identify the capacity level of the organisation for improving its recordkeeping.

A1.3.4.3.3.1.2, Design Transition Strategies

Use information about business performance, the infrastructural framework and the capacity assessment to produce transition strategies within the constraints of financial plans and the business framework.

A1.3.4.3.3.1.3, Establish RK Implementation Plan

Define scope, goals, resources and deliverables of implementation plans based on the capacity assessment, the proposed transition strategies, the financial plans and the recordkeeping performance, and develop performance criteria.

A1.3.4.3.3.1.3.1, Define RK Scope, Goals, Resources and Deliverables

Transform the transition strategy into implementation scope and goals, needed resources and deliverables, given the capacity assessment.

A1.3.4.3.3.1.3.2, Define RK Implementation Plan

Within the constraints of an organisation's capacity assessment, the financial plans and the current recordkeeping framework, define the specifications for implementing the organisation's recordkeeping function through consideration of the defined objectives and deliverables, of the evaluation of the performance of the recordkeeping function, and of updated information regarding the progress of the actual implementation.

A1.3.4.3.3.1.3.3, Develop RK Implementation Performance Criteria

Develop performance criteria for monitoring the progress of implementation of recordkeeping within the organisation and the achievement of the objectives set in the implementation plans, based on the implementation goals, resources, and identify what information should be measured and provided in implementation progress reports.

A1.3.4.3.3.2, Develop Recordkeeping Instruments

Develop metadata sets and rules, metadata encoding schemes, appraisal instruments, and education program and courses within the given business and recordkeeping frameworks.

A1.3.4.3.3.2.1, Develop RK Metadata Schemas and Rules

Define metadata requirements and rules, identify and select appropriate metadata standards and sets, and if needed construct metadata schema(s).

A1.3.4.3.3.2.1.1, Identify RK Metadata Requirements

Analyze business metadata structures, match recordkeeping needs for metadata, identify potential metadata sources, and define metadata specifications.

A1.3.4.3.3.2.1.1.1, Analyse Business Metadata Structures
Identify and analyse existing business metadata schema's that should be taken into account and can be used in developing a recordkeeping metadata framework.

A1.3.4.3.3.2.1.1.2, Match Recordkeeping Needs for Metadata
Match the recordkeeping needs for metadata with the identified business metadata sets to identify potential sources for extracting recordkeeping metadata.

A1.3.4.3.3.2.1.1.3, Establish RK Metadata Specifications
Specify metadata requirements based on recordkeeping metadata needs, the integrated recordkeeping functionality and identified metadata sources in the business processes.

A1.3.4.3.3.2.1.2, Select Appropriate Metadata Standards/Sets
Select appropriate recordkeeping metadata standards and sets from identified and available recordkeeping metadata standards and existing sets and assess their usability based upon the recordkeeping metadata requirements within the given recordkeeping framework.

A1.3.4.3.3.2.1.3, Develop Metadata Structures
Construct customised recordkeeping metadata schemas taking into account the selected and usable components from existing recordkeeping metadata schema's, based upon the identified recordkeeping metadata requirements.

A1.3.4.3.3.2.2, Develop RK Metadata Encoding Schemes
Develop classification schemas, controlled vocabularies, thesauri, and a system for unique and persistent identification as needed within the given recordkeeping framework.

A1.3.4.3.3.2.2.1, Develop Classification Schemes
Assess the metadata requirements and the integrated RM functionality to develop an activity classification scheme and an access classification scheme.

A1.3.4.3.3.2.2.2, Develop Controlled Vocabularies/Thesaurus
Based upon the metadata requirements, the classification schemes and the functional analysis develop the controlled vocabularies and thesauri that are needed.

A1.3.4.3.3.2.2.3, Establish System for Unique Identification
Identify or design and subsequently implement a system for unique and persistent identifiers (within the organisation or beyond depending on the needs) of records and/or their aggregates.

A1.3.4.3.3.2.3, Develop Retention Schedule
Based on the results of the functional analysis, the analysis of the usage of records and their aggregates, and an analysis of the relationships between business systems develop a retention schedule.

A1.3.4.3.3.2.3.1, Analyse Use of Records
Analyse information about relevant business processes based upon the functional analysis, consolidated usage information, and business need(s) for records and other identified retention requirements as laid down in the business and recordkeeping frameworks to identify primary and secondary uses of records.

A1.3.4.3.3.2.3.2, Analyse Relationships Between Business Systems
Based upon the primary and secondary uses of records, and the infrastructural framework assess the relationships between business systems in order to identify the primary sources of records.

A1.3.4.3.3.2.3.3, Compile Retention Schedule

Based upon the functional appraisal report, information about relevant relationships between systems, primary and secondary uses of records, and the recordkeeping framework compile the appropriate retention schedule.

A1.3.4.3.3.2.4, Develop RK Education Program and Courses

Based on the recordkeeping framework, customised schemas, the retention schedule and activity and access classification schemes develop an education program and courses in close relation with the recordkeeping implementation plans.

A1.3.4.3.3.3, Design Recordkeeping Infrastructure

Analyse integrated recordkeeping functionality, recordkeeping instruments, the infrastructural framework and retrieval requirements and design recordkeeping processes and identify requirements for recordkeeping in systems.

A1.3.4.3.3.3.1, Analyse Recordkeeping System Requirements

Analyse recordkeeping instruments, integrated recordkeeping functionality and retrieval requirements under the constraints of the metadata and recordkeeping framework to produce requirements for recordkeeping processes and recordkeeping in systems.

A1.3.4.3.3.3.2, Design Recordkeeping Processes

Design recordkeeping processes from recordkeeping requirements in systems based on the infrastructural framework and within the recordkeeping framework.

A1.3.4.3.3.3.3, Design RK Technical Infrastructure

Design the technical infrastructure for recordkeeping based on the design of recordkeeping processes, recordkeeping requirements for systems and the infrastructural framework given the state of technology.

A1.3.4.3.3.4, Train Staff

Use recordkeeping education program and courses, the recordkeeping instruments and information about staff and assigned responsibilities (recordkeeping framework) to train staff.

A1.3.4.3.3.5, Monitor/Evaluate RK Implementation

Based on the recordkeeping framework, the recordkeeping implementation plans, and implementation performance criteria, the results of training, and the evaluation information of recordkeeping performance monitor and evaluate the progress of recordkeeping implementation and produce progress reports.

A1.3.4.3.4, Evaluate RK Performance and Adequacy

Assess based on business and recordkeeping performance whether the way the recordkeeping framework is implemented and carried out in the organisation is still appropriate based upon the established recordkeeping performance criteria. Based on this analysis monitoring reports will be produced to inform the recordkeeping framework management function to confirm or revise the recordkeeping framework.

A1.3.4.4, Develop Infrastructural Framework

Design, develop, implement and monitor a comprehensive, administrative and operational infrastructure with integrated recordkeeping functionality and processes to effectively manage (plan, build, use and leverage) all of an organisation's strategic framework components and compile performance and progress reports.

A1.3.4.5, Establish Integrated Strategic Framework

Integrate the business and recordkeeping frameworks into one robust and functional framework that will govern all business and recordkeeping activities, systems, and processes.

A1.4, Define Business Processes and Structures

Define an organisation's business operational targets and outcomes, delegate and assign resources, develop business and workplans, and design and develop business work processes, necessary instruments and systems structures to effectively manage the organisation's resources and support its work processes.

A1.4.1, Define Business Targets and Outcomes

Develop business plans with an organisation's current and future performance goals, operational priorities, and product/service development and delivery strategies and objectives in line with the organisation's mandate and business and accountability framework, and, as necessary to continue to meet evolving business and stakeholder needs and interests, refine these targets and outcomes in response to information about ongoing business performance.

A1.4.2, Assign Resources

Based on the objectives and directives of an organisation's business plan(s), develop work plans for the allocation of the organisation's human and material resources needed to achieve the business targets and to support business work processes, instruments and systems, and, as necessary to continue to meet evolving business and stakeholder needs and interests, refine these work plans in response to ongoing business workflow and performance information.

A1.4.3, Develop Work Processes

Based on the objectives of an organisation's business work plan(s), analyse business work processes and develop workflows for operationalising the organisation's those work processes, and, as necessary to continue to meet evolving business and stakeholder needs and interests, refine the work process analysis and workflow strategy in response to business work plan revisions.

A1.4.4, Design and Develop Business Instruments and Systems

In response to the particular operational, material, technological, or other requirements related to an organisation's work plans and workflows, create, install or otherwise provide the instruments and systems needed to support the execution of an organisation's business work processes.

A1.5, Monitor/Evaluate Business Performance

Periodically assess the performance of the business processes in relation to the organisation's strategic framework and the accountability framework. Based on the monitoring, produce business performance reports to inform the organisation's appropriate management functions to confirm or revise the business strategic framework, or business processes and structures.

A2, Carry Out Business Activity

Carry out a business function, activity or perform task or a set of one or more related activities to provide services, make or deliver services, products or any other specified result.

A3, Manage Records

Manage in an efficient and systematic way the capture, receipt, maintenance, use and disposition of records, including processes for capturing and maintaining evidence of and information about business activities and transactions in the form of records

A3.1, Capture Records

Based on rules established in the recordkeeping framework the capture function identifies and brings under control the records that are created in the business activity and need to be maintained. With the capture of those records, the required metadata are also captured/extracted to ensure the authenticity, usability, integrity and reliability of the records. The capture of metadata is done every time a record or aggregation of records is used in a business process. The capture process includes the registration and classification of the records as well as, if needed, the assignment of key words, so that the records are identifiable and searchable. The valid RK instruments will guide the registration and classification. Identification and information about the performance of this function are produced for evaluation purposes.

A3.1.1, Document Records and Their Provenance

Attach to records all the contextual metadata, needed to know when, why, by whom and in what business process the record has been created and/or used, as well as the interrelationships with other records in order to ensure the authenticity, usability and reliability.

A3.1.2, Register Records

Assign a unique identifier to records and/or their aggregates in accordance with the identification and registration rules and document the registration.

A3.1.3, Classify Records

Classify identified and captured records or aggregates by assigning a classification code from the classification scheme and, if necessary, add key words for retrievability.

A3.2, Maintain Records

Following direction established in the preservation strategy as part of the recordkeeping framework for a given body of records selected for preservation, apply preservation method(s) targeted to that body of records to implement the preservation action plan for those records by maintaining the digital components of accessioned digital records, along with related information necessary to reproduce the records, certify their authenticity and enable correct interpretation of the records.

The maintain activity carries out also the disposition function, so that records are kept no longer than needed.

This maintenance activity enables the output, in response to a retrieval request, of the digital components of a record, along with information about that record, or, if the request is only for information, the requested information. The ‘maintain’ process also produces management information which is used to evaluate execution of the ingest function. The process is carried out by persons responsible for preservation, using infrastructure technology.

A3.2.1, Manage Preservation of Records

Provide overall control and co-ordination of the records preservation function via preservation action plans issued in response to information about the current state of technology, the organisation’s preservation strategy, updated storage information and ongoing performance information.

A3.2.2.1, Manage Information About Records

Collect and maintain information necessary to carry out the Preservation Action Plan in support of the overall Preservation Strategy for a body of electronic records being preserved, including information about their digital components, the archival aggregates they comprise, their authenticity, their interpretation, and the preservation activities performed on them. In carrying out actions specified in the Preservation Action Plan, information about (captured / accessioned) Electronic Records is collected when they are accessioned and is combined with Storage Information identifying the files, locations, and other relevant data about the digital components of the (captured/ accessioned) Electronic Records when they are placed in storage and subsequently when storage parameters are changed.

When a Preservation Action Plan entails any modifications to digital components, Information About those Digital Components is provided to ensure that all affected components are updated appropriately and, after the modification, Information about the Updated Digital Components is also updated.

In response to a Retrieval Request for information, Retrieved Information About a Preserved Record is provided. In response to a Retrieval Request for a record, information identifying the digital components of the record and their storage location(s) is retrieved to produce a Request for Digital Components, which is used to retrieve those components from storage; Information About those Digital Components and Retrieved Information About the Preserved Record is output to support reproduction of the record and, if needed, certification of its authenticity.

A3.2.2.1, Maintain Information About Records

Manage information about the provenance, aggregation, content, form, structure, or other essential characteristics of records and/or their digital components that are needed to preserve them over time as well as to satisfy user requests.

A3.2.2.2, Retrieve Information About Records

Output information about the provenance, aggregation, content, form, and structure of records in storage.

A3.2.2.3, Retrieve Information About Digital Components

Output technical information concerning digital components of records that is required to facilitate preservation and updating of digital components.

A3.2.3, Manage Storage of Digital Components

In accordance with the Preservation Strategy established for a body of records, and applying the Storage Method selected to implement that Strategy, place the digital components of (Accessioned) Electronic Records into storage, taking the specific steps defined in the Preservation Action Plan for these records and maintain them. In response to a Request for Digital Components, retrieve the requested components and output them.

When digital components are output for updating in accordance with a Preservation Action Plan, place the Updated Digital Components in storage and, as provided by the Action Plan, either maintain or delete the older versions of these components. Provide to the ‘Manage Information’ process Updated Storage Information about the identities, locations and other relevant parameters of stored digital components whenever components are updated or other changes, such as media refreshment, are made in storage.

A3.2.3.1, Place Digital Components in Storage

Store digital components of records, if required into one or more digital containers for preservation purposes, in accordance with the Storage Plan outlined in the Preservation Action Plan.

A3.2.3.2, Refresh Storage

Convert storage of digital components of records from one medium to another, or otherwise ensure that the storage medium remains sound, in accordance with the Storage Update Method outlined in the Preservation Action Plan.

A3.2.3.3, Monitor Storage

Supervise the operation of the storage system, the media on which the digital components of records are stored, the digital components, and the facilities where the system and components are located, in accordance with the Monitoring Method outlined in the Preservation Action Plan.

A3.2.3.4, Correct Storage Problems

Take the appropriate actions prescribed by the Problem Correction Method outlined in the Preservation Action Plan to eliminate any identified problem regarding the storage of digital components.

A3.2.3.5, Retrieve Digital Components from Storage

Output copies of retrieved digital components of records in storage, in accordance with the Retrieval Method outlined in the Preservation Action Plan, in response to requests for records that consist of those components, and, in cases where digital components are encountered that need updating, redirect them to be updated.

A3.2.4, Update Digital Components

As indicated by the Preservation Strategy established for a given body of electronic records (tunnelled to this diagram), take the steps indicated in the applicable Preservation Action Plan, applying the Method(s) for Updating Components prescribed by the strategy to update Digital Components of a Record that cannot be Preserved because of technological obsolescence, changes in Preservation Strategy, or similar factors. Examples of update processes include migration, standardisation, and transformation to persistent form. Return the Updated Digital Components to Storage, providing Information about the Updated Digital Components to the 'Manage Information' process. If the Updated Digital Components belong to a record that is the subject of a Retrieval Request, also send the components, along with related information, to the Output Electronic Record process. However, if the updating was done only to satisfy conditions of a Retrieval Request and was not required to conform to Preservation Strategy, the Updated Digital Components are sent, along with related information, to the Output Electronic Record process, but they are not sent to storage.

This process may be invoked directly when records in a transfer are being examined and it is determined that there is a need to take action to preserve a record, before the components are sent to storage.

A3.2.4.1, Migrate Digital Components to Current Formats

When the format of a type of digital components is obsolete, use a targeted preservation strategy to migrate these digital components to the chosen current format, and associate with this group of digital components documentation of the actual migration.

A3.2.4.2, Convert Digital Components to Standard Formats

When the format of a type of digital components is proprietary, use a targeted preservation strategy to convert/migrate these digital components in proprietary format to the chosen standard format.

A3.2.4.3, Transform Digital Components to Persistent Formats

When the preservation action for the digital components of a record aggregation is to transform them to persistent format, apply the method for transforming the digital components to a software and hardware independent format that has been identified as persistent.

A3.3, Facilitate Access

Governed by the access framework, support search facilities for users and, if successful, provide information about or provide access to reproduced (authentic) records or produce, if requested, a reproducible digital record; that is, the digital component(s) of the record along with instructions for producing an authentic copy of the record and information necessary to interpret the record as kept under the regime of the recordkeeping framework.

A3.3.1, Manage Access

Control and co-ordinate all actions and the authoritative, procedural, and technological competences to retrieve, represent/reproduce, read, annotate, transfer, and/or destroy stored records and define the access procedures within the given access framework.

A3.3.2, Search for Information

Based on a request for records or information about records search and retrieve all components and/or information that comply to this request and can be processed for output. In case of no hit produce a negative response.

A3.3.3, Process Requests

Verify requests for records based upon access procedures and the search results, reject requests if access is restricted or requester is not authorised and inform requester, define further specifications if needed (e.g. redaction of records), facilitate retrieval of records or their aggregates and their digital components, and account for any problems with requests in accordance with the prescribed Access Procedures.

A3.3.4, Output Records

Check correctness, integrity and completeness of retrieved records or their aggregates, their digital components, and/or information about requested records, reconstitute requested records and either present them as reproduced records (with a certificate of authenticity if required) with accompanying information, or package the outputs as reproducible records.

A3.3.5, Monitor Usage and Access

Monitor access to records and their components and the information about them based upon the usage information, collect and compile statistical information about usage, analyse any failures or mistakes related to access and report this to the manage access function.

A3.4, Collect Performance Information

Synthesise and compile reports on the performance of the capture, maintain and facilitate access functions based on information continuously collected from these functions in order to inform the manage recordkeeping framework function. These reports may contain information about the applicability of policies, rules and methods, deviations from policies/rules, malfunctioning of systems, as well as suggestions for improvement. Other reports will be made with consolidated information about usage of records or aggregations.

Business-driven Recordkeeping Model Arrow Definitions

Access classification scheme
A systematic classification of categories of access rights and restrictions used for the purpose of controlling access to stored records; aggregates and/or information about records.

Access framework
A comprehensive set of administrative and operational rules, procedures and methods that will guide and control the use of records/aggregates and the information about them, and the outreach within the given business context for the identified and targeted communities and users, and taking into account their requirements for access and use of the records.

Access procedures
A set of regularised, administrative steps taken in procuring, granting or denying access to records/aggregates and/or information about records and to administrative and operational functions within an organisation's recordkeeping system in a certain business context.

Access rights
Indicators of which users have what degree of access to which resources and to what administrative and operational functions within an organisation's recordkeeping system.

Accountability framework
A comprehensive, integrated, operational and administrative set of policies, procedures and rules that provides for: (1) consistent management of an organisation's various (external) stakeholders, (2) continuous monitoring and assessment of compliance to the accountability framework within the organisation, (3) continuous monitoring and assessment of the impact of new business processes, legislation, policies, juridical and social requirements or other relationships on the attribute of accountability within the organisation, and (4) modification of the organisation's accountability framework design, as necessary, in response to these monitoring and assessment activities.

Accountability reports
Reports that provide information on the state and performance of the management of the records continuum in the related business context.

Activity classification scheme
A plan based on the systematic identification and arrangement of an organisation's business activities into categories that will govern the arrangement of records and/or aggregates.

Aggregation information
Information about rules for aggregating records, and the agreed levels of aggregation.

Analysis of legal and juridical requirements
An analysis of an organisation's legal and juridical obligations in relation to its business and the effect with respect to the recordkeeping requirements needed to meet those obligations.

Appraisal policy
A set of cohesive and coherent policies, procedures, rules, standards, guidelines, criteria and methods for determining what records should be created/captured, how long they should be kept and how they should be preserved, based on, among other things, consideration of the organisation's recordkeeping risk analysis as well as evaluation of record context, value (i.e., current and future uses), and preservation feasibility.

Appraisal policy evaluation reports
An evaluation of the suitability of the current appraisal policy based upon the recordkeeping performance information and the consolidated records usage information. It contains an assessment on how the current appraisal policy suits the needs of the organisation.

Assigned RK responsibilities
Formal, documented indications of which (groups of) users are assigned what responsibilities with respect to the administrative and operational functions with respect to an organisation's recordkeeping.

Business framework
A comprehensive, integrated, operational and administrative structure that will guide, and govern an organisation's business processes, people, systems, operations and projects in line with the organisation's overall mandate.

Business performance reports
Reports that state the performance of the various business activities of an organisation based on particular benchmark criteria provided in the integrated strategic framework.

Business plans
Operational plans that spell out an organisation's expected course of action for a specified period in relation to one or more business activities and, for example, services and products, the market, the industry, management policies, marketing policies, production needs and available resources, and usually including an analysis of risks and uncertainties and how to mitigate them.

Business systems and instruments
All systems, instruments and tools needed to support and conduct the business activities.

Capacity assessment
Assessment of the current situation in the organisation with respect to creating, managing, and preserving records and determining what needs to be done to improve it to the next level as identified and agreed. The capacity assessment includes assessing the level of awareness and understanding of management and staff, the available technical infrastructure, the level of expertise of the Recordkeeping staff, the available resources, ...

Classification scheme + identifier
A logical plan based on subjects or business activities used for classifying records/ aggregates.

Communication/ transactions
Messages or other information exchanges between persons and/or systems in the course of a business activity.

Consolidated records usage information
Aggregated information regarding: (1) the history of successful, unsuccessful and rejected access requests for records (and/or information), including data about the nature of the requests, about which records (and/or information) were successfully accessed when and by whom, about which access requests were unsuccessful and why, and about which access requests were rejected and why, (2) the purposes for which the records (and/or information) were requested, (3) the retrieval questions asked, (4) the frequency of requests for records (and/or information), and (5) problems encountered in fulfilling requests or retrieving requested records (and/or information).

Contextual analysis
The integration of an organisation's mandate, mission, business function and activities, history, corporate culture, strengths, weaknesses, market position, and legal, social and corporate responsibilities.

Controlled vocabulary/thesaurus
A managed set of purposefully delimited and standardised terms, phrases and concepts used by an organisation to control the values of a metadata element.

Corporate culture
An organisation's values, beliefs, business principles, traditions, ways of operating, and internal work environment.

Customised schemas
Recordkeeping metadata schemas customised according to targeted recordkeeping metadata requirements and based upon selected recordkeeping metadata standards.

Defined RK objectives and deliverables
An accounting of the overall scope, performance goals and objectives, and expected output of an organisation's recordkeeping function, and the resources required to support that function.

Design of RK processes
An organisation's formal plan outlining the related, structured activities that are part of each discrete recordkeeping process, the rules governing them and their workflow.

Digital components of records
One or more binary (digital) components in a certain format that are required to reconstitute a record.

Digital components that need updating
Digital components of maintained records that cannot be reconstituted or presented in accordance with current maintenance strategies applicable to those records.

Education program and courses
An education program and courses in understanding the need for recordkeeping as well in skills in using and applying recordkeeping instruments such as metadata schemas for record, an activity classification scheme, retentions schedules, in line with the responsibilities of the staff involved.

Essential characteristics of record types
The intrinsic and extrinsic characteristics of records, as well as the technological characteristics of their digital components, identified as critical for reflecting the intent of the record in the business activity. They regard form, structure, content, context and sometimes behaviour.

Evaluation of RK performance
Consolidated information derived from the evaluation of performance gathered from the recordkeeping processes under control of the given recordkeeping framework. The assessment is done based on the performance criteria.

Existing business metadata structures
The metadata schemas and encoding schemes in use in the business activities and systems.

Existing sets of functional RK requirements
Sets of functional requirements for recordkeeping that are available as standard or made by other organisations or institutions

Existing standards or sets of functional requirements to be used
Choice of (parts of) existing sets of functional requirements for recordkeeping that may fulfil the needs of the organisation in this area within the given infrastructural framework.

Experiment designs
Designs for carrying out experiments to get a better insight in the usability of available preservation strategies for certain types of digital components. They will be properly documented.

External mandate
External directive, law or decision for assigning the responsibility for a business function or set of business activities.

Facilities
Physical space, systems, resources and technical infrastructure needed to enable the business.

Financial information
Information collected and required by an organisation to support its financial decisions or to meet its financial requirements.

Financial plans
A budgetary planning document reflecting the way an organisation plans to acquire or earn and use its financial resources in a given year.

Functional analysis
An analysis of an organisation's business functions and supporting business processes, that identifies how they are interrelated including, in particular, the key relationships between the constituent operational and administrative elements (tasks and transactions).

Functional requirements for RK
The requirements for recordkeeping functionality in systems.

Guidelines and directives for appraisal
The set of guidelines and rules that will govern the appraisal policy and processes.

Guidelines for organising records
A set of guidelines and procedures for structuring records and/or aggregates, for developing classification schemes based upon business activities, and for identifying aggregation layers.

Identified metadata sources
Metadata sources that could serve as a source for extracting recordkeeping metadata.

Identified records
Records or their aggregates to which the contextual metadata (e.g., persons, actions and dates) have been attached.

Identified set of RK functional requirements
The set of functional recordkeeping requirements derived from preservation and appraisal policies, access framework, synthesised records requirements and evaluation information of recordkeeping performance, needed to serve the organisations needs for managing records.

Indexes
Tools that facilitate efficient and effective location records and/or records aggregates suited to a particular inquiry or business purpose. Indexes may be supported by recordkeeping Instruments, such as controlled vocabularies and thesauri.

Information about business performance
Information about the way business activities are carried out in accordance with the strategic framework and the business plans, as well as information about deviations and needs and other relevant evaluation information.

Information about destroyed records
Formal information documenting the destruction of records, including information about the quantity and characteristics of records that have been destroyed, and the reason why.

Information about digital components
Technical information concerning digital components of records that is required to facilitate updating of the digital components.

Information about records
Information about the provenance, aggregation, content, structure, form, format or other characteristics of records and/or aggregates in storage.

Information about related records
Information about other records participating in the same business process or activity.

Information about relevant business processes
Description of those business processes that should and will produce/create and need records. This description includes an analysis of the constituting steps, transactions, responsibilities, constraints, workflow, and the related required records.

Information about updated digital components
Information about changes that have been made to digital components of records in the process of updating them, about any problems that occurred in the process, the dates of the updating, and the persons responsible for the updating.

Information identifying digital components of a requested record
Technical composition and location information concerning digital components of a requested record that is necessary to reconstitute the record from its digital components.

Infrastructural framework
A comprehensive, integrated, operational and administrative set of policies, standards, procedures and guidelines for managing all the ‘hard’ and ‘soft’ structural elements that support the human and organisational capabilities required to effectively plan, build, use and leverage all of an organisation’s integrated framework components.

Integrated RK functionality
The consolidated set of functional requirements necessary to support the recordkeeping framework and the underlying recordkeeping processes.

Integrated strategic framework
A comprehensive administrative and operational set of policies, standards, procedures and guidelines that includes the integration of an organisation’s strategic risk management, business and recordkeeping frameworks, and that will govern and guide the organisation’s business activities as well as the support of business and recordkeeping applications and processes.

Internal mandates
The authority derived from external mandates and within the given strategic framework, invested by management or a corporate board or subsidiary to perform specific functions.

Internal requests for records or information
Requests from internal users to consult or receive records, their aggregates, or information about records.

IPR policy
An authoritative set of cohesive and coherent policy, rules, guidelines, and methods for protecting the intellectual property rights of the organisation or other stakeholders from unauthorised access or use.

Juridical requirements
The current laws and regulations that govern the organisation and its business activities and recordkeeping, and/or social, professional, or sectoral/domain specific requirements that influence the need for records.

Keywords
Key terms that have been attributed to a record or aggregate.

Maintained information about digital components
Information about the technical composition and location of digital components that is necessary to maintain/update, store and retrieve the components, and to reconstitute the records from their components.

Maintained information about records
Information about the provenance, aggregation, content, form, structure, and behaviour of records, and/or about the administrative and technical information about them.

Mandate
The legal basis or the intention to achieve a certain goal. This can be translated into different mission statements for carrying out a certain business depending on the interpretation, the identified needs at a certain moment in time, and the available resources.

Mandate drivers
The most basic, core incentives for carrying out business activities.

Market analysis
An evaluation of the business environment, the potential markets in relation to the organisation's main (envisaged) products or services, potential customers, possible competitors, and of what its strengths and weaknesses are, especially in relation to its customers and competitors.

Market information
Information produced or collected by an organisation about the organisation's products and/or services, its customers, and its competitors for the purpose of supporting the organisation's economic, operational and/or competitive success.

Monitoring method
The preservation oversight method stipulated in the Preservation Action Plan for determining whether a storage system is properly maintained and functioning or whether storage media are intact and free from problems that would interfere with reading the data written on the media.

Outgoing communication
Messages and other documents communicated with external parties in relation to the business activities, as well as the final result(s) of carrying out business activities.

Performance analysis
An assessment, based on consideration of business performance reports and other formal and informal data, of the actual performance or functionality of an organisation's business framework at any point in time, relative to the desired performance or functionality of the business framework as specified in the business framework plan.

Performance information
Information about the performance of recordkeeping activities, including registration of exceptions, errors, inadequacy or deviations of established rules and methods.

Preservation action plans

Set of rules, procedures, methods, and technical requirements governing the specific preservation actions to be taken for the classes of digital components covered by the Preservation Policy, and which also indicate the time or conditions when such actions should be taken.

Preservation policy

The authoritative set of coherent policies, standards, guidelines, and criteria for maintaining and preserving records, their aggregates and their related metadata as well as their constituent digital components, as long as required according to the retention policy. These policies and standards include guidelines and criteria for maintaining digital components, and for reconstituting and reproducing records in authentic form. The policy is taking into account the evaluation of the recordkeeping framework, performance information, prior preservation policies and the state of technology.

Preservation policy evaluation reports

Definition 1: An evaluation of the suitability of the current preservation policy based upon the recordkeeping performance information.

Definition 2: Assessment information on how the current preservation policy suits the needs of the organization.

Preservation principles

The core and authoritative instructions governing the process of preserving records that are used to help guide the preservation policy.

Preservation requirements

The requirements identified to preserve the record types, aggregations, associated metadata and digital components, for as long as they need to be preserved.

Primary and secondary use

The value of records both with respect to their ability to serve the purposes of the business and the organisation and to their ability to serve as sources of information for persons and organisations other than the business and the organisation (e.g. freedom of information, collective memory).

Privacy policy

An organisation's authoritative set of cohesive and coherent rules and guidelines, and methods for protecting the information on (human) subjects contained in the organisation's records from unauthorised access.

Problem correction method

The strategy stipulated in the preservation action plan for correcting storage and/or access problems of a specified type.

Processed requests

Requests for information about records or aggregates or for records/aggregates and/or their digital components that have been accepted as eligible for processing.

Product or service

Any product or service that is the result of a business activity and falls under the mandate of the organisation.

Record to be created

Those records deemed necessary and/or desirable to support business needs and obligations as determined based on the appraisal policy.

Records
Information created, received, and maintained as evidence and information by an organisation or person, in pursuance of legal obligations or in the transaction of business (ISO 15489-1:2001).

Recovered file
A physical or logical file to which the problem correction method prescribed in the preservation action plan has been successfully applied after one or more storage problems affecting the file were found.

Refreshed file
A physical or logical file that have been copied from an older storage medium or system to a newer one using the storage update method prescribed in the preservation action plan.

Registered and classified records
Records that have been checked, registered and classified according to the rules and classification scheme and that need to be stored.

Registered records
Records that have been assigned an unique identifier and the necessary registration metadata, according to the registration rules.

Registration information
Information about the records that have been captured, including an unique identification number, registration date, and other registration information deemed necessary.

Registration rules
A formal set of rules specifying how records need to registered and what metadata need to be captured and what procedures for assigning a unique identifier to each record need to be followed.

Rejected requests
Requests that do not meet the criteria or requirements for access. Reasons for rejection may include restrictions on records or information about them, incomplete requests.

Relevant relationships between systems
Analysis of the relationships between an organisation's various business (administrative and operational) systems, and the business functions, processes and transactions they support, necessary to develop an appropriate retention schedule.

Reproduced/reproducible records and information about them
Reproduced Electronic Record. An authentic representation or other version of a record reconstituted from its digital components, along with information supporting the interpretation of the record.

Request for digital components
An instruction to retrieve the digital components of one or more records or of one specific type (format).

Requested digital components/packages
Digital components of one or more records retrieved from storage based on a request either as individual entities or wrapped in metadata and with instructions.

Requested records
Records requested either by internal users in support of business activities, or by external users for support of their own activities.

Requests for records or information

Requests from either internal or external users to consult or receive records, their aggregates, or information about records.

Requirements for presenting records/aggregates

An accounting of the contextual and provenancial characteristics of records and records aggregates necessary for ensuring that records are presented properly ordered, identified and documented with respect to their relationships with each other. As a result the records/ aggregates can be properly interpreted and used.

Requirements for RK in system(s)

The functional requirements for recordkeeping in systems necessary to adequately support an organisation's Recordkeeping processes within the constraints of available technology.

Retention requirements

Requirements that from a business perspective determine how long records should be preserved based on the identified needs and risks.

Retention schedule

An overview of how long records and/or aggregates need to be kept based upon a functional classification business records in an organisation, including authorised disposition instructions for those records.

Retrieval method

The method identified in a preservation action plan for retrieving information about records/aggregates and/or retrieving the digital components in response to requests for records or information about or contained in records.

Retrieval requirements

The set of requirements that should support the retrieval of records/aggregates and information about records/aggregates and their digital components based upon the usage information and patterns of users either within or outside the organisation.

Rights and responsibilities

The set of legal, moral and/or ethical expectations of, and obligations to, any and all affected parties with respect to the impact of such expectations and obligations on an organisation's activities, employees, partners, clients, etc.

Risk analysis

Report on an analysis of the risks that may exist based on an assessment of existing internal and external requirements and frameworks, the business needs and ensuring the continuity of an organisation. This risk analysis identifies the consequences for the creation, management, maintenance of and access to records (and provides options how to take care of them)..

Risk management framework

A set of cohesive and coherent policies, procedures, rules, standards, guidelines, criteria and methods based on identified risks, the probability of these risks occurring, (possible) measures to mitigate them, including the assignment of responsibilities.

This framework will be based on an analysis of the rights and responsibilities, stakeholder interests, mandate drivers and market analyses within the constraints of society, the organisation's accountability framework and its corporate culture.

RK framework
A set of cohesive and coherent policies, procedures, rules, standards, guidelines, criteria and methods that will guide, manage and align an organisation's recordkeeping functions with the organisation's juridical requirements and its business and risk management frameworks.

RK HRM instruments
Description of recordkeeping competencies and of recordkeeping job functions as well as information about education/training of staff in recordkeeping skills.

RK implementation performance criteria
The set of criteria used to measure the progress and success of the implementation of the recordkeeping program (or parts of it) within the organisation.

RK implementation plans
Plans stating the scope and goals of implementation of recordkeeping in the organisation, including the required resources and deliverables

RK implementation plans & progress reports
Plans stating the scope and goals of recordkeeping implementation, including the required resources and deliverables, as well as reports detailing the current state of implementation progress.

RK implementation progress reports
Reports detailing the current state of progress in implementing recordkeeping in relation to the goals and deliverables.

RK instruments
Instruments that will support the creation and management of records/aggregates, such as classification scheme(s), retention schedule, controlled vocabularies, access classification scheme.

RK metadata requirements
Metadata specifications based on recordkeeping metadata needs within the given business context, the integrated recordkeeping functionality and identified metadata sources.

RK performance criteria
Operational criteria that enable the assessment of the recordkeeping activities, processes, procedures, and systems in relation to the requirements and the goals set in the recordkeeping framework

RK performance reports
Reports with periodically compiled information about the ability of the individual components of the recordkeeping function to meet the performance criteria.

RK principles, standards and methods
The principles, standards and methods governing the management of records within the Recordkeeping profession.

RK process and system requirements
The set of requirements for recordkeeping processes and for supporting systems and technical infrastructure, needed to fulfill the identified needs of the organisation for the creation, maintenance and disposal of records.

RK risk analysis
An evaluation of the business context and related legal, business, organisational and societal requirements to identify risks to which the recordkeeping function needs to respond, their probability and how to mitigate them.

RK security policy
An organisation's authoritative set of cohesive and coherent rules/guidelines, criteria and methods for establishing, updating and enforcing access rights to records based on (re)evaluation of the organisation's recordkeeping infrastructure framework and recordkeeping system performance, in concert with the constraints of the organisation's appraisal policy and risk management framework.

RK theory and methods
The existing theoretical concepts, principles, standards, tools, and approaches in the recordkeeping domain

Sample records
A set of records well-described and randomly retrieved from storage, which, along with preservation performance information, are used to support/evaluate preservation strategies, tests. There may be different sets of different composition (e.g., by type, or by complexity)

Search results
Information provided to the user based on his/her search question that can be used determine what records should be retrieved or to refine a search question.

Selected metadata components
Metadata standards and sets selected from recordkeeping framework and integrated RM functionality.

Societal influence
Social, moral and ethical standards and expectations of (external) communities.

Specifications for authentic records
Information identifying and describing/specifying what essential characteristics of records determine the authenticity of records seen from the perspective of the business activity.

Staff
An organisation's personnel who have been assigned tasks.

Stakeholder interest information
Information about the interests of an individual or group in the performance of an organisation, its capability to be profitable, in delivering the intended results/products, and maintaining the viability of the organisation's products and/or services.

Stakeholder interests
The interests of an individual or group with respect to the success of an organisation in delivering intended results and maintaining the viability of the organisation's products and/or services.

State of technology
The state of the art of the information technology with respect to its ability to satisfy recordkeeping requirements, the state of the underlying computer science with respect to its ability to develop relevant capabilities not within the state of the technology, and the existence and prevalence of applicable standards.

Storage plan
Set of rules, procedures and technical requirements governing storage of the digital components of records. Also part of the Preservation Action Plan that specifies files and directories (and their relative and absolute locations) into which one or more digital components of records are placed in the storage system.

Storage problem
A problem with storage media, storage formats, digital components, a storage system or facility that could impact on access to, and the continued preservation of, stored records.

Storage update method
A method used to ensure that stored digital components are completely and correctly brought forward when any component of a storage subsystem is changed or when digital components are moved or migrated to newer storage media.

Stored digital file
Digital files placed in a storage system on digital media.

Synthesised records requirements
The coherent set of requirements from a business perspective that include the characteristics critical for authenticity for each type of records, rules for metadata, and the rules for organising the records.

System for unique identifiers
A systematic way of identification that is documented and consistently applied and enables the assignment and persistently linking of a unique, unambiguous, and permanent code (potentially any combination of numeric and alphabetical values) to every record and/or aggregation.

Targeted preservation methods
For specific record formats, methods for overcoming technological obsolescence of file formats. Includes methods such as emulation, conversion to current format, conversion of proprietary formats to standard formats, and conversion to self-describing file formats.

Technical infrastructure
The planned configuration of an organisation's computing systems (hardware and software), associated communications and network systems, as well as any services that may be needed to operate and maintain such systems.

Trained staff
All employees who have been trained to perform specific Recordkeeping activities within an organisation. This includes also staff that is not exclusively working within the recordkeeping function.

Transition strategies
A description of the strategies and related activities necessary to move the organisation from the current level of recordkeeping to the identified next level, based upon the capacity assessment. This includes criteria to assess whether the organisation has reached the next level.

Unique identifier
The code with which a record or aggregate uniquely can be identified. The domain of uniqueness will be at least the organisation. The code should be persistent through time.

Updated digital components
An updated digital component is a component that has been modified under a preservation action plan.

Updated storage information
Information indicating a change in the location of a digital component in storage, the occurrence of a storage problem, the action taken to correct a storage problem, the results of such action, or the copying of digital files from older to new storage media.

Usage information

Information about the need for and the use of records, the questions asked, the frequency of use, problems in retrieval.

Work plans

Schedules, charts, graphs or other documents that summarises clearly the various components of particular business processes or activities and how they fit together, while outlining the significant responsibilities and duties of individuals or entities with respect to those processes or activities.

Work process analysis

The precise mapping of the sequence of steps or actions required to produce a business outcome that complies with the organisation's functions, its systems and rules. It may be extended to describe the derivation of the organisational procedures and rules from the socio-legal context in which the organisation is located.

Workflow

A descriptive and analytical account of one or more work processes, including the sequence of inherent steps, required to achieve a defined outcome.

It describes the transfer of documents, records, information and/or tasks from one participant (human or machine) to another for action, according to a set of procedural rules and within the constraints of the current state of technology.

Business-driven Recordkeeping Model

Definitions of the classes occurring in the UML class diagrams

11 January 2008 Version
by Hans Hofman

> Note: The following list of definitions of the entities used in the UML class diagrams of the BDR model (see Narrative associated with the Consultation Draft of the model) are derived from many different sources and may be subject to comments.

Access action
The action of finding, using, or retrieving information, usually subject to rules and conditions. *(NAA Recordkeeping Glossary)*

(Business) activity
Activities are the major tasks performed by an organisation to accomplish each of its functions. An activity is identified by the name it is given and its scope note. The scope of the activity encompasses all the transactions that take place in relation to it. Depending upon the nature of the transactions involved, an activity may be performed in relation to one function, or it may be performed in relation to many functions. *(NAA Recordkeeping Glossary)*

> *Business activity is used as a broad term, not restricted to commercial activity, but including public administration, non-profit and other activities.*

Agent

> **Responsible agent**

Individual, workgroup or organization responsible for or involved in record creation, capture and/or records management processes (*ISO 23081-1:2006*).

Analysis
A systematic approach to problem solving. Complex problems are made simpler by separating them into more understandable elements. This involves the identification of purposes and facts, the statement of defensible assumptions, and the formulation of conclusions.
www.ojp.usdoj.gov/BJA/evaluation/glossary/glossary_a.htm

> **Risk analysis**
> **Market analysis**

Archive(s)
1. Records that are appraised as having archival value. *(NAA Recordkeeping Glossary)*
2. A level of aggregation of records generated by an individual or organisational body.

Behaviour
The set of dynamic and interactive characteristics of a digital record that are necessary for using it in a way that reflects the original intent and thus its authenticity.

Business rules
Set of the operations, definitions and constraints that apply to an organization in achieving its goals.

> For example a business rule might state that no credit check is to be performed on return customers. http://en.wikipedia.org/wiki/Business_rules

Capture
A deliberate action that results in the registration of a record into a recordkeeping system. For certain business activities this functionality maybe built into computer systems so that the capture of records is concurrent with the creation of records. (*NAA recordkeeping glossary*)

Code of practice
A set of written rules that state operating requirements for specified activities. www3.gov.ab.ca/env/air/Info/definitions.html

Code of ethics
An organized group of ethical behaviour guidelines, which govern the day-to-day activities of a profession or organization. www.peakagents.ca/glossary/c10.htm

Computer file
A set of data with the same file format.

Cultural values
The accumulated habits, attitudes, and beliefs of a group of people *(or social group)* that define for them their general behaviour and way of life; the total set of learned activities of a people. www.geographic.org/glossary.html

Customer relations
The interaction between an organisation (or individual) and one or more persons who expect, request, buy something from it.

Disposition
Any action that changes the circumstances of a record or removes a record from its usual setting. Disposal can include destruction, damage, alteration, or transfer of custody or ownership of records. The National Archives of Australia authorises disposal of Commonwealth records for the purposes of the *Archives Act 1983*. Also called disposal, usually in the Australian context. (*NAA recordkeeping glossary*)

Encoding scheme
Controlled list of all the acceptable values in natural language and/or as a syntax-encoded text string designed for machine processing (*ISO 23081-1:2006*).

Extraction action
A preferably automated activity that extracts, based on a cross-walk between two metadata schema's, (meta)data from one environment (mostly system) into another.

NB. In case it is not automated the tool has to be informed manually what metadata should be extracted.

Form
The extrinsic elements of a record that reflect the lay-out, structure and if needed the behaviour.

File
(Noun) An organised unit of documents accumulated during current use and kept together because they deal with the same subject, activity or transaction. (*NAA recordkeeping glossary*)

Framework
A system of rules, ideas or principles that is used to plan or decide something.
www3.gov.ab.ca/env/air/Info/definitions.html

Risk management framework
Strategic framework
Business framework
Infrastructural framework
Accountability framework
Financial framework

Function
Functions represent the major responsibilities that are managed by the organisation to fulfil its goals. They are high-level aggregates of the organisation's activities. *(NAA Recordkeeping Glossary)*

Jurisdiction
The extent of authority: 1) of a court over a certain matter or person; 2) of a government organization. (adapted from www.websiteupgrades.ca/glossary/free/J.shtml)

Legislation
A law including acts of parliament and other general legal rules.
www.apheda.org.au/campaigns/burma_schools_kit/resources/1074040257_16812.html

Mandate
Identifies and provides information about the instrument that imposes a requirement to make and keep a record or group of records.

Internal mandate
External mandate

Mandates can be internal or external. Internal mandates include policy, administrative instructions, business decisions or authorisations. External mandates include laws, regulations, standards or statements of best practice etc that incorporate requirements to make and keep records. *(NSW RKMS)*

Market
A public place where goods and services are traded, purchased and sold.
www.valic.com/valic2003/aigvalic.nsf/contents/edu_glossary-m

or

1. The interaction between supply and demand to determine the market price and corresponding quantity bought and sold.
2. The determination of economic allocations by decentralized, voluntary interactions among those who wish to buy and sell, responding to freely determined market prices.
www-personal.umich.edu/~alandear/glossary/m.html

Metadata
Data describing the context, content and structure of records and their management through time (*ISO 15489-1:2001, 3.12*). As such, metadata are structured or semi-structured information that enables the creation, registration, classification, access, preservation and disposition of records through time and within and across domains.

> **Structural metadata:** metadata documenting the record structure
> **Technical metadata:** metadata documenting the technical environment in which the record has been originally generated as well as the technical environment necessary to reproduce it
> **Contextual metadata:** metadata documenting the provenance of the record or any aggregation

Plan
A proposed or intended method of getting from one set of circumstances to another. They are often used to move from the present situation, towards the achievement of one or more objectives or goals. http://en.wikipedia.org/wiki/Plan

> **Risk management plan**
> **Work plan**
> **Budget plan**
> **Resource plan**

Preservation action
The processes and operations involved in ensuring the physical, technical and/or intellectual survival of authentic records through time.

Preservation encompasses environmental control, security, creation, storage, handling, and disaster planning for records in all formats, including digital records. *(NAA Recordkeeping Glossary)*

Record
Information created, received, and maintained as evidence and information by an organization or person, in pursuance of legal obligations or in the transaction of business (*ISO 15489:2001, clause 3.15*).

Record instance
The description of how a record can be expressed or reproduced without losing its authenticity, reliability and integrity, using one or more specified computer files.

Record characteristics
The coherent set of essential aspects that given the business context in which a record (or aggregate) is created or used, make that record authentic, reliable, and usable. They regard the structural, behavioural and form aspects of a record.

Requirements
Business requirements
RK requirements

Rights
Entitlements assured by law, custom, or property.
[IPR, copy rights, privacy]

Risk
The possibility of suffering loss, harm, danger, damage
Business risk
RK risk

Schema
RK metadata schema
Business metadata schema
Logical plan showing the relationships between metadata elements, normally through establishing rules for the use and management of metadata specifically as regards the semantics, the syntax and the optionality *(obligation level)* of values (*ISO 23081-1:2006*).

Series
A level of aggregation of records, above the case file.
NB. May be named differently in different jurisdictions.

Social structure
[nation, community, family]
A system of social relations.

Social structure does not concerns itself with people—individuals forming the society or their social organisations, neither does it study who are the people/organisation forming it, or what is the ultimate goal of their relations. Social structure deals rather with the very structure of their relations—how are they organized in a pattern of relationships.
http://en.wikipedia.org/wiki/Social_structure

Technical environment
The platform of hard- and software needed to reproduce or render authentic records in digital form.

Transaction
The smallest unit of business activity. Uses of records are themselves transactions.
See also **Function** and **Activity**.

Sources: Adapted from Standards Australia, AS 4390, Part 1, Clause 4.27; Standards Australia, AS ISO 15489, Part 2, Clause 4.2.2.2.
RK transaction
Business transaction

Transformation
It holds information on the process of performing a migration. It consists of a collection of transformation units (which in turn hold information on which files have been migrated into what).

Transformation unit
A logical unit that will be migrated together as one unit. This could be a single file (e.g., a Word document) or a large set of files (e.g., an entire database of multiple files). It contains two sets of files: the Pre-Transformation File Set (the set of files that are to be or have been migrated) and the Post-Transformation File Set (the set of files created as a result of this transformation).

APPENDIX 16

Overview of the Records Continuum Concept

[Extracted and adapted from: Xiaomi An, "An Integrated Approach to Records Management," *Information Management Journal* July/August (2003): 24–30.]

Overview of the Records Continuum Concept[1]

The following text was extracted and adapted from:

> Xiaomi An, "An Integrated Approach to Records Management," *Information Management Journal* July/August (2003): 24–30.

As defined in Australian Standard 4390, the records continuum is "...a consistent and coherent regime of management processes from the time of the creation of records (and before creation, in the design of recordkeeping systems) through to the preservation and use of records as archives."[2]

The earliest view of the continuum concept came from Australian national archivist Ian Maclean in the 1950s. He said records managers were the true archivists, and that archival science should be directed toward studying the characteristics of recorded information, recordkeeping systems, and classification processes. His view promoted the search for continuity between archives and records management.

The records continuum as a model concept was formulated in the 1990s by Australian archival theorist Frank Upward based on four principles:

5. A concept of "record" inclusive of records of continuing value (archives) stresses their uses for transactional, evidentiary, and memory purposes, and unifies approaches to archiving/recordkeeping, whether records are kept for a split second or a millennium.
6. There is a focus on records as logical rather than physical entities, regardless of whether they are in paper or electronic form.
7. Institutionalization of the recordkeeping profession's role requires a particular emphasis on the need to integrate recordkeeping into business and societal processes and purposes.
8. Archival science is the foundation for organizing knowledge about recordkeeping. Such knowledge is revisable but can be structured and explored in terms of the operation of principles for action in the past, the present, and the future.[3]

In her book *Yesterday, Today and Tomorrow: A Continuum Responsibility*, Sue McKemmish writes: "The model provides a graphical tool for framing issues about the relationship between records managers and archivists, past, present, and future, and for thinking strategically about working collaboratively and building partnerships with other stakeholders."[4]

In *Records Management: A Guide to Corporate Recordkeeping*, Jay Kennedy and Cherry Schauder explain the four dimensions that Upward used in his concept of the continuum model:

9. At level one, the model identifies accountable acts and creates reliable evidence of such acts by capturing records of related/supporting transactions. Records of business

---

[1] See also http://www.sims.monash.edu.au/research/rcrg/.

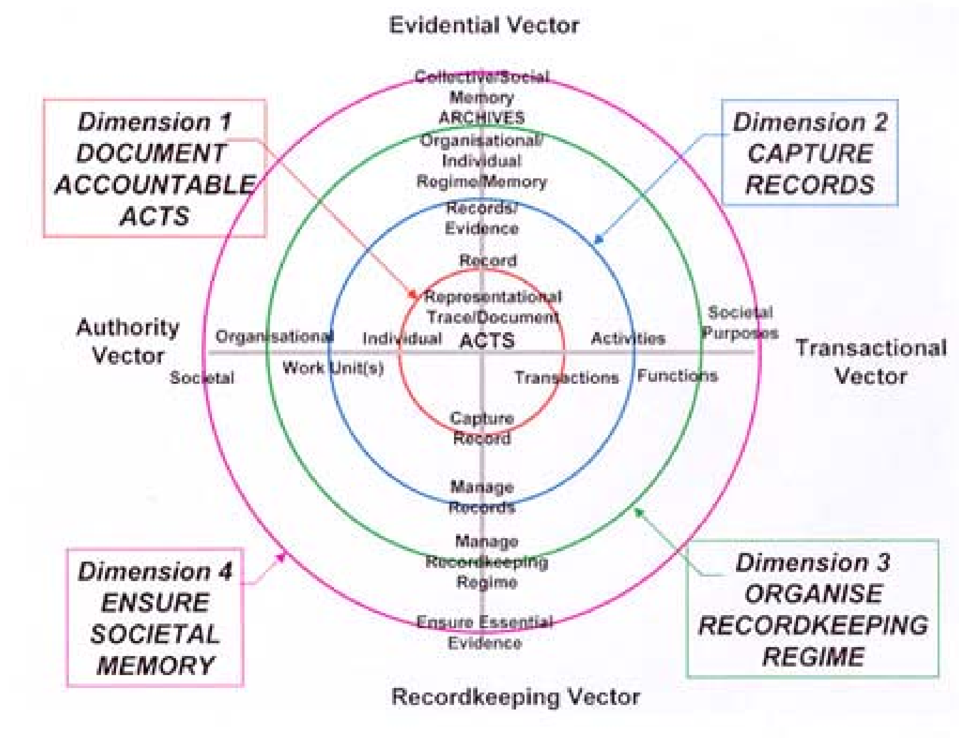
[2] Standards Australia, *Australian Standard 4390, Records Management* (Homebush, New South Wales: Standards Australia, 1996).

[3] Frank Upward, "In Search of the Continuum: Ian Maclean's 'Australian Experience' Essays on Recordkeeping," in S. McKemmish and M. Piggot, eds. *The Records Continuum: Ian Maclean and Australian Archives First Fifty Years* (Sydney: Ancora Press in association with Australian Archives, 1994). Online reprint available at http://www.sims.monash.edu.au/research/rcrg/publications/fuptrc.html.

[4] See Sue McKemmish (1997), "Yesterday, Today and Tomorrow: A Continuum Responsibility," in *Proceedings of the Records Management Association of Australia 14th National Convention*, RMAA Perth, 15-17 September 1997.

activities are created as part of business communication processes within the organization (e.g., through e-mail, document management software, or other software applications).

10. At level two, recordkeeping systems manage "families" of transactions and records series documenting processes at the work-unit or single-function scope of complexity. Records that have been created or received in an organization are tagged with metadata, including how they link to other records.
11. At level three, a seamless recordkeeping scheme embraces the multiple systems and families of records that serve the entire documentary needs (i.e., business, regulatory, and cultural/educational/historical) of a single juridical entity. Records become part of a formal system of storage and retrieval that constitutes the organization's corporate memory.
12. At level four, a collaborative recordkeeping establishment under the guidance of a suitably empowered public recordkeeping authority serves the needs of the total society, its constituent functions, and the entities that carry them out. The recordkeeping establishment serves the documentary needs of many entities within its jurisdiction and ensures the accountability and the cultural memory of the society as a whole. Records required for purposes of societal accountability (e.g., by corporate law) or other forms of collective memory become part of wider archival systems that comprise records from a range of organizations.[5]



[5] Jay Kennedy and Cherryl Schauder, *Records Management: A Guide to Corporate Recordkeeping*, 2nd edition (South Melbourne: Longman, 1998).

In the article "The Records Continuum Model in Context and Its Implications for Archival Practice," Sarah Flynn explains that the records continuum model is significant because it

- broadens the interpretation of records and recordkeeping systems offered by the lifecycle model. Such broadening is helpful, given the variety of contexts in which archivists and records managers operate and in which archives and records are used;
- reminds us that records (including archives) are created and maintained for use as a result of business and administrative functions and processes, rather than as ends in themselves; and
- emphasizes cooperation beyond the walls of repositories, especially between the closely related, if occasionally estranged, professions of archives administration and records management—a cooperation that is more important than ever in the contemporary climate of outsourcing and cross-sector working.[6]

In the article "Life Cycle Versus Continuum: What Is the Difference?" Peter Marshall states that the records continuum's primary focus is the multiple purposes of records.[7] It aims for the development of recordkeeping systems that capture, manage, and maintain records with sound evidential characteristics for as long as the records are of value to the organization, any successor, or society. It promotes the integration of recordkeeping into the organization's business systems and processes.

According to McKemmish, the best-practice mechanism behind the records continuum model uses an integrated approach for managing records and archives. Records managers and archivists are brought together under an integrated recordkeeping framework with the same goal: to guarantee the reliability, authenticity, and completeness of records. The framework provides common understanding, consistent standards, unified best-practice criteria, and interdisciplinary approaches and collaborations in recordkeeping and archiving processes for both paper and digital worlds. It provides sustainable recordkeeping to connect the past to the present and the present to the future. It can coherently exist in a broader dynamic, changeable context that can be influenced by legal, political, administrative, social, commercial, technological, cultural, and historical variables across time and space. The integrated recordkeeping framework would:

- facilitate provenance
- underpin accountability
- constitute memory
- construct identity
- provide authoritative sources of value-added information

The continuum's purpose-oriented, systems approach to records management fundamentally changes the role of recordkeeping. Instead of being reactive, managing records after they have been created, recordkeeping becomes proactive. In partnership with other stakeholders, identifying records of organization activities that need to be retained, then implementing business systems designed with built-in recordkeeping capability ensures capturing records of evidential quality as they are created. Built-in capture and assessment mean that records of value are created in the first place whenever electronic systems are used for business transactions. With appropriate metadata to ensure that they are accurate, complete, reliable, and usable, these records have the necessary attributes of content, context, and structure to act as evidence of business activity. And, Marshall notes, knowing from the outset which digital records must be

[6] Sarah J. A. Flynn (2001), "The Records Continuum Model in Context and its Implications for Archival Practice," *Journal of the Society of Archivists* 22(1): 79–93.
[7] Peter Marshall (2000), "Life Cycle Versus Continuum: What is the Difference?" *Informaa Quarterly* 16(2): 20–25.

kept for the longer term means such records can be migrated across systems as hardware and software upgrades occur.

The mechanisms of best practice behind the records continuum model are ideal for integrating records and archives management because the records continuum focuses on:

- similarities rather than differences
- qualities and quantities rather than quantities alone
- positive and cohesive ways of thinking rather than disparate or passive ways
- integrated policy making rather than fragmented frameworks
- integrated control of policy implementation rather than separate control
- integrated rather than disparate approaches to problem solving
- meeting customers' needs through collaboration rather than by duplication and overlap

These arguments highlight the records continuum model's importance as a best-practice model for managing digital records when the aim is to improve responsiveness, increase efficiency and satisfy users' requirements.

APPENDIX 17

Metadata Schema Analysis Questions

Metadata Schema Analysis Questions

QID	Question	IP	ISO	CV Type	Controlled Vocab
General Schema Information					
100	Is the schema intended for recordkeeping purposes?			Single CV Select (without Element List)	Yes#No
101	Whether or not the schema is intended for recordkeeping purposes, indicate which recordkeeping entities can be described. Select all that apply.			Multiple CV Select (without Element List)	Agents#Business Processes#Mandates#Records#Records Management Processes#Relationships#
102	Are the elements of the schema categorized according to a particular scheme?			Single CV Select (without Element List)	Yes#No
102	If the elements in the schema are categorized according to a particular scheme, indicate the categories used to classify the elements.			Open End	
103	In general, which of the following views of records management metadata can the schema address? Select all that apply.			Multiple CV Select (without Element List)	Business Perspective#Records Management Perspective#Use Perspective#
104	Schemas may be classified as single object schemas or multiple object schemas. What is the schema type of the schema?			Single CV Select (without Element List)	Single Object Schema#Multiple Object Schema
Record Metadata Questions					
200	IDENTIFICATION OF RECORD AGGREGATION LEVELS			Dummy	
201	How does the schema describe a record object?			Multiple CV Select (without Element List)	Element Set#Element(s)#Other#Not Applicable
202	How does the schema describe record aggregations?			Multiple CV Select (without Element List)	Element Set#Element(s)#Other#Not Applicable
203	How does the schema describe recordkeeping systems or corporate archives?			Multiple CV Select (without Element List)	Element Set#Element(s)#Other#Not Applicable
204	How does the schema describe collective archives?			Multiple CV Select (without Element List)	Element Set#Element(s)#Other#Not Applicable

205	GENERAL RECORD METADATA			Dummy	
206	Does the schema contain elements to identify date related information for a record? If so, indicate the applicable element(s).	A1.a	9.2.1 a	Single CV Select with Element List	Yes#No
206	Indicate which of the following date types the schema describes. Select all that apply.	A1.a.iii		Multiple CV Select (without Element List)	Chronological Date#Creation Date#Received Date#Archival Date#Transmission Date#
207	Does the schema contain elements to identify time related information for a record? If so, indicate the applicable element(s).	A1.a	9.2.1 a	Single CV Select with Element List	Yes#No
207	Indicate which of the following time types the schema describes. Select all that apply.	A1.a.iii		Multiple CV Select (without Element List)	Creation Time#Compiled Time#Received Time#Archival Time#Transmission Time#
210	STRUCTURAL METADATA			Dummy	
211	Does the schema contain elements that directly identify record structure? If so, indicate the applicable element(s).		9.2.1 c	Single CV Select with Element List	Yes#No
212	Does the schema contain elements that directly identify record form? If so, indicate the applicable element(s).		9.2.1 d	Single CV Select with Element List	Yes#No
213	Does the schema contain elements that directly identify the chemical and physical properties of the record? If so, indicate the applicable element(s).	A3;A4	9.2.1 e	Single CV Select with Element List	Yes#No
214	Does the schema contain elements that directly identify the technical characteristics and dependencies of a record? If so, indicate the applicable element(s).	A4	9.2.1 f	Single CV Select with Element List	Yes#No
215	Does the schema contain elements that directly identify the technical requirements to render or reproduce records? If so, indicate the applicable element(s).	B1.c	9.2.1 h	Single CV Select with Element List	Yes#No
216	RECORD ACCESS			Dummy	
217	Does the schema contain elements that directly identify record aggregation level? If so, indicate the applicable element(s).	A1.a.iv; B3	9.2.1 m;9.2.3.1 a	Single CV Select with Element List	Yes#No
218	Does the schema contain elements that directly identify record location information (whether physical or logical)? If so, indicate the applicable element(s).	A2; A7; A8; B1.a; B3	9.2.3.1 c;9.6.1 h	Single CV Select with Element List	Yes#No
219	Does the schema contain elements that directly identify record title? If so, indicate the applicable element(s).	A1.a.iv	9.2.3.1 e;9.6.1 h	Single CV Select with Element List	Yes#No
220	Does the schema contain elements that directly identify record subject classification? If so, indicate the applicable element(s).	A1.a.iv	9.2.3.1 e; 9.6.1 h	Single CV Select with Element List	Yes#No
221	Does the schema contain elements that directly identify record descriptive keywords? If so, indicate the applicable element(s).	A1.a.iv	9.2.3.1 e;9.6.1 h	Single CV Select with Element List	Yes#No
222	Does the schema contain elements that directly identify a record abstract? If so, indicate the applicable element(s).	A1.a.iv	9.2.3.1 e;9.6.1 h	Single CV Select with Element List	Yes#No

223	Does the schema contain elements that directly identify the language of a record? If so, indicate the applicable element(s).	A1.a.iv	9.2.3.1 e;9.6.1 h	Single CV Select with Element List	Yes#No
224	Does the schema contain elements that directly identify record indexing? If so, indicate the applicable element(s).	A1.a.iv	9.2.3.1 e;9.6.1 h	Single CV Select with Element List	Yes#No
225	RECORD SECURITY			Dummy	
226	Does the schema contain elements that directly identify record and record aggregation access restrictions? If so, indicate the applicable element(s).	B1.b	9.2.4.1 a	Single CV Select with Element List	Yes#No
227	Does the schema contain elements that directly identify business process access restrictions? If so, indicate the applicable element(s).	B1.b	9.2.4.1 a	Single CV Select with Element List	Yes#No
228	Does the schema contain elements that directly identify agent access restrictions? If so, indicate the applicable element(s).	B1.b	9.2.4.1 a	Single CV Select with Element List	Yes#No
229	Does the schema contain elements that directly identify time limitations on record access restrictions? If so, indicate the applicable element(s).	B1.b	9.2.4.1 c	Single CV Select with Element List	Yes#No
230	Does the schema contain elements that directly identify time limitations on business process access restrictions? If so, indicate the applicable element(s).	B1.b	9.2.4.1 c	Single CV Select with Element List	Yes#No
231	Does the schema contain elements that directly identify time limitations on agent access restrictions? If so, indicate the applicable element(s).	B1.b	9.2.4.1 c;9.6.1 e	Single CV Select with Element List	Yes#No
232	RELATIONSHIPS TO THE RECORD			Dummy	
233	Does the schema document the relationship between the record and the business transaction or activity that generated it? If so, indicate the applicable element(s).		9.2.1 l	Single CV Select with Element List	Yes#No
234	TBD Does the schema capture the relationships between the data and the format element(s) comprising the record? If so, indicate the applicable element(s).		9.2.1 g	Single CV Select with Element List	Yes#No

Agent Metadata Questions

300	IDENTIFICATION OF AGENT GROUPS			Dummy	
301	How does the schema identify a person or actor?			Multiple CV Select (without Element List)	Element Set#Element(s)#Other#Not Applicable
302	How does the schema identify organizational unit and/or work group?			Multiple CV Select (without Element List)	Element Set#Element(s)#Other#Not Applicable
303	How does the schema identify organizations or corporate bodies?			Multiple CV Select (without Element List)	Element Set#Element(s)#Other#Not Applicable

304	How does the schema identify social institutions?			Multiple CV Select (without Element List)	Element Set#Element(s)#Other#Not Applicable
305	GENERAL AGENT METADATA			Dummy	
306	Does the schema contain elements that directly identify a creator of records? If so, indicate the applicable element(s).	A1.a.i	9.4.1 a;9.2.1 b	Single CV Select with Element List	Yes#No
307	Does the schema contain elements that directly identify agents involved in records management processes? If so, indicate the applicable element(s).	A1.a.i	9.4.1 b	Single CV Select with Element List	Yes#No
308	Does the schema contain elements that directly identify agents involved in records transactions (addressees, originators, etc.)? If so, indicate the applicable element(s).	A1.a.i	9.5.1 c	Single CV Select with Element List	Yes#No
309	AGENT SECURITY	A2		Dummy	
310	Does the schema contain elements that directly identify agent authorizations for records management processes? If so, indicate the applicable element(s).	B1.b; B2.a	9.4.1 b	Single CV Select with Element List	Yes#No
311	Does the schema contain elements that directly identify access privileges for authorized agents to create, use, modify, destroy, etc. records? If so, indicate the applicable element(s).	B1.b	9.4.1 c;9.6.1 d	Single CV Select with Element List	Yes#No

Business Process Metadata Questions

400	IDENTIFICATION OF BUSINESS PROCESS CATEGORIZATIONS			Dummy	
401	How does the schema identify business transactions?			Multiple CV Select (without Element List)	Element Set#Element(s)#Other#Not Applicable
402	How does the schema identify business activities?			Multiple CV Select (without Element List)	Element Set#Element(s)#Other#Not Applicable
403	How does the schema identify business functions?			Multiple CV Select (without Element List)	Element Set#Element(s)#Other#Not Applicable
404	How does the schema identify ambient functions?			Multiple CV Select (without Element List)	Element Set#Element(s)#Other#Not Applicable
405	GENERAL BUSINESS PROCESS METADATA			Dummy	
406	Does the schema contain elements that directly identify business processes? If so, indicate the applicable element(s).		9.5.2 a	Single CV Select with Element List	Yes#No
407	Does the schema contain elements that directly identify the security and access rules for business processes and transactions? If so, indicate the applicable element(s).	B1.b	9.5.1 d	Single CV Select with Element List	Yes#No

408	Does the schema contain elements that directly identify business process classification schemes? If so, indicate the applicable element(s).		9.2.3.1 f; 9.5.1 f; 9.6.1 g	Single CV Select with Element List	Yes#No
409	Does the schema contain elements that directly identify the date of a transaction? If so, indicate the applicable element(s).	A1.a.iii	9.5.1 h	Single CV Select with Element List	Yes#No
410	Does the schema contain elements that directly identify the time of a transaction? If so, indicate the applicable element(s).	A1.a.iii	9.5.1 h	Single CV Select with Element List	Yes#No
411	RELATIONSHIPS TO BUSINESS PROCESS			Dummy	
412	TBD Does the schema document the relationship between business processes and the records? If so, indicate the applicable element(s).		9.5.1 b	Single CV Select with Element List	Yes#No
413	Does the schema document the relationship between business processes and agents? If so, indicate the applicable element(s).		9.5.1 b	Single CV Select with Element List	Yes#No

Mandate Metadata Questions

500	INTERNAL BUSINESS RULES & SYSTEM CONTROLS			Dummy	
501	Does the schema contain elements to identify metadata schemes or schemas used in organization business systems? If so, indicate the applicable element(s).		9.3.1 a	Single CV Select with Element List	Yes#No
503	Does the schema contain elements to identify rules or controls that regulate record creation? If so, indicate the applicable element(s).		9.3.1 b	Single CV Select with Element List	Yes#No
504	Does the schema contain elements to identify rules or controls that regulate records management? If so, indicate the applicable element(s).	B1.b	9.3.1 b	Single CV Select with Element List	Yes#No
505	Does the schema contain elements to identify rules or controls that regulate record access? If so, indicate the applicable element(s).	B1.b	9.3.1 e	Single CV Select with Element List	Yes#No
506	Does the schema contain elements to identify rules or controls that regulate the rights to records? If so, indicate the applicable element(s).	B1.b	9.3.1 e	Single CV Select with Element List	Yes#No
507	Does the schema contain elements to identify rules or controls for records management operations? If so, indicate the applicable element(s).	B1.b	9.3.1 d	Single CV Select with Element List	Yes#No
508	Does the schema contain elements to identify rules or controls for metadata creation? If so, indicate the applicable element(s).		9.3.1 c	Single CV Select with Element List	Yes#No
509	Does the schema contain elements to identify rules or controls for metadata management? If so, indicate the applicable element(s).	B1.b	9.3.1 c	Single CV Select with Element List	Yes#No
510	EXTERNAL MANDATES AND REGULATIONS			Dummy	

511	Does the schema contain elements to identify mandates or regulatory requirements for record creation? If so, indicate the applicable element(s).		9.3.1 f	Single CV Select with Element List	Yes#No
512	Does the schema contain elements to identify mandates or regulatory requirements for records management? If so, indicate the applicable element(s).		9.3.1 f	Single CV Select with Element List	Yes#No
513	Does the schema contain elements to identify mandates or regulatory requirements for record retention? If so, indicate the applicable element(s).		9.3.1 g	Single CV Select with Element List	Yes#No
514	Does the schema contain elements to identify mandates or regulatory requirements for record security? If so, indicate the applicable element(s).		9.3.1 g	Single CV Select with Element List	Yes#No
515	Does the schema contain elements to identify mandates or regulatory requirements for record destruction? If so, indicate the applicable element(s).		9.3.1 g	Single CV Select with Element List	Yes#No
516	RELATIONSHIPS TO INTERNAL AND EXTERNAL MANDATES			Dummy	
517	Does the schema document the relationship between external mandates and regulations and the records? If so, indicate the applicable element(s).		9.3.1h	Single CV Select with Element List	Yes#No
518	Does the schema document the relationship between external mandates and regulations and internal records management operations? If so, indicate the applicable element(s).		9.3.1h	Single CV Select with Element List	Yes#No

RK Process Metadata Questions

600	GENERAL RECORDKEEPING METADATA			Dummy	
601	Does the schema contain elements to identify retention information applied to a record? If so, indicate the applicable element(s).	B1.a	9.6.1 b	Single CV Select with Element List	Yes#No
602	Does the schema contain elements to identify disposition authority schemes? If so, indicate the applicable element(s).		9.6.1	Single CV Select with Element List	Yes#No
603	Does the schema contain elements to identify security classification schemes? If so, indicate the applicable element(s).		9.6.1	Single CV Select with Element List	Yes#No
604	Does the schema contain elements to identify a record's registration into a record system? If so, indicate the applicable element(s).		9.6.1	Single CV Select with Element List	Yes#No
605	LONG-TERM PRESERVATION			Dummy	
606	Does the schema contain elements to document conversion of records from one format or medium to another? If so, indicate the applicable element(s).		9.6.2 a;9.2.1 i;9.2.2	Single CV Select with Element List	Yes#No

607	Does the schema contain elements to document the agent responsible for record conversion? If so, indicate the applicable element(s).	B2.a	9.6.2 a;9.2.2	Single CV Select with Element List	Yes#No
608	Does the schema contain an element(s) to document the date and time of record conversion? If so, indicate the applicable element(s).	B2.a	9.6.2 a;9.2.2	Single CV Select with Element List	Yes#No
609	Does the schema contain an element(s) to document any changes created by the record conversion? If so, indicate the applicable element(s).	B2.c	9.6.2 a;9.2.2	Single CV Select with Element List	Yes#No
611	(TBD)?? Relationship between the records acquired from the creator and the copies produced by the preserver.	B2.b		Single CV Select with Element List	Yes#No
612	(TBD) Does the schema contain an element(s) to document the reason for record conversion? If so, indicate the applicable element(s).		9.6.2 a;9.2.2	Single CV Select with Element List	Yes#No
616	Does the schema contain elements to document migration of records from one hardware or software system to another? If so, indicate the applicable element(s).	B2.c		Single CV Select with Element List	Yes#No
617	Does the schema contain elements to document the agent responsible for record migration? If so, indicate the applicable element(s).			Single CV Select with Element List	Yes#No
618	Does the schema contain an element(s) to document the date and time of record migration? If so, indicate the applicable element(s).			Single CV Select with Element List	Yes#No
619	Does the schema contain an element(s) to document any changes created by the record migration? If so, indicate the applicable element(s).			Single CV Select with Element List	Yes#No
620	(TBD) Does the schema contain an element(s) to document the reason for record migration? If so, indicate the applicable element(s).			Single CV Select with Element List	Yes#No
621	Does the schema contain elements to document other procedures to counteract media fragility and technological obsolescence ? If so, indicate the applicable element(s).	A4		Single CV Select with Element List	Yes#No
624	RECORD BACKUP			Dummy	
625	Does the schema contain elements to document regular record back-ups performed on records? If so, indicate the applicable element(s).	A3		Single CV Select with Element List	Yes#No
626	Does the schema contain elements to document the agent responsible for record backups? If so, indicate the applicable element(s).	A3		Single CV Select with Element List	Yes#No
627	Does the schema contain an element(s) to document the date and time of record backup? If so, indicate the applicable element(s).	A3		Single CV Select with Element List	Yes#No
628	Does the schema contain an element(s) to document any changes created by the record backup? If so, indicate the applicable element(s).	A3		Single CV Select with Element List	Yes#No

629	(TBD) Does the schema contain an element(s) to document the reason for record backup? If so, indicate the applicable element(s).			Single CV Select with Element List	Yes#No
635	RECORD REPRODUCTION			Dummy	
636	Does the schema contain elements to identify record reproduction activities (other than backups) for a record? If so, indicate the applicable element(s).			Single CV Select with Element List	Yes#No
637	Does the schema contain an element(s) to indicate the agent responsible for record reproduction? If so, indicate the applicable element(s).	B2.a		Single CV Select with Element List	Yes#No
638	Does the schema contain an element(s) to indicate the date and time of record reproduction? If so, indicate the applicable element(s).	B2.a		Single CV Select with Element List	Yes#No
639	Does the schema contain an element(s) to document the relationship between records acquired from the creator and copies produced by the preserver? If so, indicate the applicable element(s).	B2.b		Single CV Select with Element List	Yes#No
640	Does the schema contain an element(s) to identify any changes to the record due to copy procedures? If so, indicate the applicable element(s).	B2.c		Single CV Select with Element List	Yes#No
641	Does the schema contain an element(s) to document changes and provide such information both to the preservers and users? If so, indicate the applicable element(s).	B2.d		Single CV Select with Element List	Yes#No
650	RECORD AUTHENTICATION			Dummy	
651	Does the schema identify authentication declarations for records or groups of records? If so, indicate the applicable element(s).	A6	9.6.1 c	Single CV Select with Element List	Yes#No
652	Does the schema identify rules under which authentication declarations are utilized? If so, indicate the applicable element(s).	A6	9.6.1 c	Single CV Select with Element List	Yes#No
653	Does the schema identify agents providing authentication declarations? If so, indicate the applicable element(s).	A6	9.6.1 c	Single CV Select with Element List	Yes#No
660	RECORD TRACKING			Dummy	
661	Does the schema contain elements to identify record access? If so, indicate the applicable element(s).	B1.b		Single CV Select with Element List	Yes#No
662	Does the schema contain elements to identify an agent who accessed the record? If so, indicate the applicable element(s).	B1.b		Single CV Select with Element List	Yes#No
663	Does the schema contain elements to identify the date a record was accessed? If so, indicate the applicable element(s).	B1.b		Single CV Select with Element List	Yes#No
664	Does the schema contain elements to identify the action taken on record? If so, indicate the applicable element(s).	B1.b		Single CV Select with Element List	Yes#No

665	Does the schema contain elements to identify the physical transfer of records from one location to another? If so, indicate the applicable element(s).	B1.a		Single CV Select with Element List	Yes#No
666	Does the schema contain elements to identify the agent responsible for the transfer of records from one physical location to another? If so, indicate the applicable element(s).	B1.a		Single CV Select with Element List	Yes#No
667	Does the schema contain elements to identify the date and time of record transfer? If so, indicate the applicable element(s).	B1.a		Single CV Select with Element List	Yes#No
668	Does the schema contain elements to identify the transfer location name? If so, indicate the applicable element(s).	B1.a		Single CV Select with Element List	Yes#No
670	Does the schema contain elements to identify the custodial transfer of records from one agent to another? If so, indicate the applicable element(s).	B1.a		Single CV Select with Element List	Yes#No
680	Does the schema capture changes to the classification of a record over time? If so, indicate the applicable element(s).		9.2.2	Single CV Select with Element List	Yes#No
681	AGENT TRACKING			Dummy	
682	Does the schema capture changes to personnel (including records creators, managers, users, etc.) over time?		9.2.3.2 b	Single CV Select with Element List	Yes#No
683	Does the schema capture changes to the agents' roles (including security classifications) over time? If so, indicate the applicable element(s).		9.4.2	Single CV Select with Element List	Yes#No
684	BUSINESS PROCESS TRACKING			Dummy	
685	TBD: Does the schema capture changes to records-producing business processes over time? If so, indicate the applicable element(s).		9.5.2;9.2.3.2 a;9.2.3.2 g	Single CV Select with Element List	Yes#No
686	Does the schema capture changes to business process classification schemes over time? If so, indicate the applicable element(s).		9.5.2	Single CV Select with Element List	Yes#No
687	MANDATE TRACKING			Dummy	
688	Does the schema capture changes to business rules and system controls over time? If so, indicate the applicable element(s).		9.3.2	Single CV Select with Element List	Yes#No
689	Does the schema capture changes to external mandates and regulatory requirements over time? If so, indicate the applicable element(s).		9.3.2	Single CV Select with Element List	Yes#No
690	METADATA RECORD TRACKING			Dummy	
691	Does the metadata schema contain elements to manage metadata about the metadata record ? If so, indicate the applicable element(s).		8.3.9.2	Single CV Select with Element List	Yes#No
691	Indicate what metadata record information the schema describes. Select all that apply.		8.3.9.2;8.3.8	Multiple CV Select (without Element List)	Date and time of metadata creation or alteration#Date and time of metadata alteration#Agent responsible for metadata creation or alteration#Activity undertaken to metadata record#Metadata version number#Security and

				access restrictions for metadata#
697	RELATIONSHIPS TO THE RECORDKEEPING PROCESS		Dummy	
699	Does the schema document the relationship between the record, the agent and recordkeeping processes the agent performs on the record? If so, indicate the applicable element(s).	9.6.1 i	Single CV Select with Element List	Yes#No

APPENDIX 18

Case Study Data Relating to Metadata

Case Study Data Relating to Metadata

Focus 1. Artistic Activities

General information regarding metadata

CS01, Arbo Cyber, théâtre (?)

The report states that no descriptive schemas and metadata are employed. However, records are classified by date of the performance (not by their date/time of digitization) to which they are linked. Individual practices are used to relate to the functional and technological needs of the *Ludosynthese*. However, the report reveals that if Arbo Cyber, theatre (?) decides to enter digital information, these properties will be limited to the programs' capabilities.

CS02, Performance Artist Stelarc

The report states that Stelarc has no recognized system of organization from an archival point of view with regard to his digital materials. Instead, the materials are arranged according to Stelarc's performance and publicity needs. No documented processes or procedures are used to identify, retrieve or access his digital materials. Although some records access and modification restrictions are in place, these do not appear to be formally documented. In effect, there are few, if any, formal recordkeeping practices and no metadata are consciously or intentionally recorded.

CS03, HorizonZero/ZeroHorizon Online Magazine & Media Database

The report states that the organization of the files pertaining to each issue of *HorizonZero* is ad hoc and is generally organized by the issue for which they were created. These files are accessible through a shared space that can be navigated using tracking software that organizes the posting into threads.

These tracker entries are saved using an archival function implemented in the tracker software (Mantis 0.18.0A4).

CS09(01), Altair4 di Roma

The report states that there is neither a recordkeeping system nor metadata schemas; however, Altair4 uses the "Where is it" program to reorganize and retrieve digital entities. To use them, it is necessary to know the filename, path and approximate date of production.

CS09(02), National Film Board

The report states that all work done using the computer as an intermediary [...] is kept on the server system and is related to a given project by the project number (assigned before a production is given the go-ahead; once a production is approved, it is given a unique production number).

CS09(03), Commercial Film Studio

The report states that only those digital entities that are archived have metadata. The standards used are Dublin Core, the Thesaurus for Graphic Materials I & II and AACR2.

CS09(04), WGBH Boston

It is important to note that the current production entity investigated during the case study consisted of a mixed analogue/digital production system. At the time of the case study, the creator was in the process of converting to a digital asset management (DAM) system, while at the same time maintaining its collection of analogue film, tapes and audio content that dates back to the 1950s. Catalogue records for these materials are kept in a FileMaker Pro 7.0

database designed and developed in-house. The DAM system is an Artesia TEAMS product that has been customized by WGBH.

CS10, Danube Exodus

File naming is largely ad hoc and some individuals develop their own system. Therefore, there is no formal recordkeeping system; furthermore, there is no system to track the changes, actions or transactions to the digital files.

CS13, *Obsessed Again...*

The report implicitly states that no metadata schemas or standards are employed. There is no formal recordkeeping system. All digital entities are stored on computer disks, which remain in the possession of the composer. These entities are only identified through the assignment of a semi-descriptive filename.

CS15, *Waking Dream*

The reports states that metadata is not consciously captured. The digital entities are kept in simple directories and are not entered in any sophisticated recordkeeping system. Professor Fels wrote the code used in *Waking Dream* and maintains it on his computer. Thus, retrieval and access of these digital entities is dependent on whether or not the computer in question contains the necessary application.

Metadata information in the 23 questions:

4d. How are the digital entities identified (e.g., is there a [persistent] unique identifier)?

CS01, Arbo Cyber, théâtre (?)

Arbo Cyber, theatre (?) does not make use of a persistent or unique identifier for electronic records, but they do use a naming convention. This was referred to during the interviews as the “nomenclature”: it makes use of a strict set of punctuation and spelling rules and relies on signifying and representative values[24]. This abbreviation code is very important in the *Ludosynthese*, as it indicates location within the site.

CS02, Performance Artist Stelarc

The digital entities are identified under project titles, event series and biographical content on the Web site.

CS03, *HorizonZero/*ZeroHorizon Online Magazine & Media Database

The digital entities are identified by naming conventions that are ad hoc, though some staff members have evolved consistent naming conventions for their own work.

CS09(01), Altair4 di Roma

These conventions comprise the folder with project name/file object name/number of version and the last version file object name/final version.

CS09(02), National Film Board

No information provided.

CS09(03), Commercial Film Studio

Strict naming conventions are used to identify the digital entities, and all those having a role in manipulating the file are required to adhere to these conventions. Among other elements, the name of the file contains information on the sequence, the scene, the name of the object as well as numerical information to identify the version. The sequence of information in the file name is:/studio/title/sequence/scene/object/version.

Interpretation of this information is as follows: "Studio" refers to the name of the studio that owns the artwork, since occasionally artwork is outsourced to another studio or a subsidiary. "Title" refers to the working title of the film being produced. "Sequence" and "Scene" refer respectively to these parts of the film (in the parlance of the studio studied, "scene" is the equivalent of "shot"). "Object" refers to the particular piece of artwork in hand. Finally, a version number is added to identify the precise iteration of the file. Sometimes in PODS (a proprietary system) or at the story stage, there is also an abbreviation for information such as the sequence date and the name of the artist. There has been some attempt to develop a consistent taxonomy. Specific terms to describe each object in development are selected in the brainstorming stage by the production team. Thus there is agreement by committee on the naming conventions to be used for each production. These, however, do not extend from one production to another.

CS09(4), WGBH Boston

Current: Yes, and the unique identifier links the catalogue red in the log with the original footage. The original footage and logs follows naming conventions that link them together and to the final production. Please see question 4(f).

DAM: Same as above.

CS10, *Danube Exodus*

No alternative attempt to apply persistent unique identifiers was noted. Most files were organized in folders whose directory structure seemed to follow the intellectual conceptualization of the project.

CS13, *Obsessed Again...*

The report states that the format of each digital file is dictated by the specifications of the individual software programs with which they were created. The NoteWriter, Max/MSP and Editor/Librarian files are proprietary, binary formats and, as such, their specifications are unreleased. The MIDI files used by the Max/MSP patches are standard text files following the MIDI specification.

CS15, *Waking Dream*

The report states that the digital entities are uniquely identified with file names and, when changes have been made, with version numbers.

18b. From what applications do the recordkeeping system(s) inherit or capture all digital entities and the related metadata (e.g., e-mail, tracking systems, workflow system, office system, databases, etc.)?

CS01, Arbo Cyber, théâtre (?)

This question does not really apply to Arbo Cyber, theatre (?), but it can be said that the documents are influenced by the programs used by the artists, such as Photoshop, Illustrator or Flash. However, the properties gained through these programs have no real significance and therefore cannot be seen to have any real value for the recordkeeping system.

CS02, Performance Artist Stelarc

The applications that Stelarc captures their digital entities and related metadata are from the following, the mail system, Web-driven database operated by Web host, Internet networks, public databases functioning as sources for data mining and conversion into performance images.

CS03, HorizonZero/ZeroHorizon Online Magazine & Media Database

The report states that the recordkeeping system is not an RMA; the documents are "captured" by transferring them from individual hard drives to the shared server space. Metadata are attached to those documents (once again, not automatically) that are subsequently transferred to the ZeroHorizon database.

CS09(01), Altair4 di Roma

Not applicable.

CS09(02), National Film Board

The records in the recordkeeping system come primarily from office systems such as Microsoft Office, as well as from various graphics systems (for photography and posters, for example).

CS09(03), Commercial Film Studio

Another database, built on FileMaker Pro and called ArchiveWorks, is used for tracking physical pieces of artwork that are not digital.

CS09(04), WGBH Boston

Current: Productions stand alone FileMaker databases feed into the Archives database.

DAM: Same as above and through direct user input.

CS10, *Danube Exodus*

None of the subjects have a formal or automated recordkeeping system, though all have some process by which records are kept. There is therefore no system in place to track changes, actions or transactions to digital files, beyond renaming by individuals and such strategies, and, as far as can be ascertained, none of the subjects employ any kind of digital or media asset management system that could perform similar functions. (It has not been possible to confirm this with C3.) All the subjects stated that they attempted to keep all relevant files, despite only really being concerned about the fate of work files and any secondary files that would allow them to remain functional. What constituted relevant or important files was largely left to the discretion of whatever individual was regarded as responsible for the project; for instance, the Project Manager at the Labyrinth Project.

CS13, *Obsessed Again...*

None.

CS15, *Waking Dream*

Not applicable.

18d. Does the recordkeeping system provide ready access to all relevant digital entities and related metadata?

CS01, Arbo Cyber, théâtre (?)

The report states that access is not direct, because the preservation strategy involves transferring records and placing them on external storage devices. Furthermore, Arbo controls their own entities without any need for particular measures of control.

CS02, Performance Artist Stelarc

Yes. Links are also present to make collaborators' Web sites and other relevant internet locations accessible. If general links become obsolete, the webmaster will keep them on the Web site as dead links. If important links become obsolete, new links will be set up to make that information accessible.

CS03, HorizonZero/ZeroHorizon Online Magazine & Media Database
Yes.
CS09(01), Altair4 di Roma
Not applicable.
CS09(03), Commercial Film Studio
Yes, access is maintained for all relevant digital entities and their metadata. Everything in the system that can be opened can be downloaded.
CS09(04), WGBH Boston
Current: No, the analogue/digital hybrid nature makes access cumbersome, though possible.
DAM: The fully digital nature of the recordkeeping system allows for greatly improved access, as well as the implementation of automatic standard language applications and thesaurus capability.
CS10, *Danube Exodus*
The report does not explicitly state how it provides access to the digital entities.
CS13, *Obsessed Again...*
Again, no system exists, but Dr. Hamel currently has ready access to all relevant digital entities.
CS15, *Waking Dream*
Not applicable.

18e. Does the recordkeeping system document all actions/transactions that take place in the system re: the digital entities? If so, what are the metadata captured?

CS01, Arbo Cyber, théâtre (?)
The lack of a true recordkeeping system makes it difficult to apply this question. The entities are saved on external storage devices; thus, it is impossible to modify them or for the system to document these modifications.
CS02, Performance Artist Stelarc
No, the Web master does not keep a record of specific updates to the Web site. The report states that the metadata are unknown.
CS03, HorizonZero/ZeroHorizon Online Magazine & Media Database
The report states there are no recordkeeping system.
CS09(01), Altair4 di Roma
Not applicable.
CS09(02), National Film Board
If different versions of digital entities are created by the animator, these must have a separate identification in order that they be retrievable. However, it is not known what metadata are captured as the NFB’s Synchrone system (an intranet comprised of an integration of multiple databases created through in-house software developments) used is unique to the National Film Board and the subject was not queried during the interview process.
CS09(03), Commercial Film Studio
No, for the moment only the check-in and check-out transactions are documented. Some transactions modify a record’s metadata, but these are not documented at present.
CS09(04), WGBH Boston
Current: Partially. Use of tapes is tracked in a FileMaker database, but re-use of shots is not tracked. DAM: Yes, each use will be noted along with versioning.

CS10, *Danube Exodus*
No.
CS13, Obsessed Again...
No such documentation exists.
CS15, *Waking Dream*
No metadata is consciously captured.

22. What descriptive or other metadata schema or standard are currently being used in the creation, maintenance, use and preservation of the recordkeeping system or environment being studied?

CS01, Arbo Cyber, théâtre (?)
The report states that FLA files in Flash allow for notes in a "grey-zone" that are inaccessible to users. They are used as memory aids, and no specific data is required. Furthermore, the notes only deal with content. These "grey-zones" also fail to capture information concerning the records themselves. The informant also did not see the use in identifying metadata. The informant had no knowledge of the information that can be captured in digital images. The only data attached to these images was that created automatically by the computer at the moment of creating and saving files.
CS02, Performance Artist Stelarc
This is unknown.
CS03, HorizonZero/ZeroHorizon Online Magazine & Media Database
No descriptive or metadata schema are consistently used for the records of *HorizonZero* pertaining to the production of each issue. There are naming conventions that describe the content of some records, but most records can be identified only by their context in the filing system.
CS09(01), Altair4 di Roma
There are no standards for activity of a creative nature. Since Altair4 uses no recordkeeping system, no reference is made to standards of description and/or indexing.
CS09(02), National Film Board
The NFB is introducing the use of MPEG-7 and MPEG-21 as standards for encoding content and rights about films. These are being introduced to simplify commercialization.
CS09(03), Commercial Film Studio
There are no standards for creation of the assets in the workflow pipeline. However, the archivist has introduced standards for description and indexing which cover those assets that make it to the archive. These include the Categories for the Description of Works of Art (CDWA), the Dublin Core (DC), the Thesaurus for Graphic Materials I: Subject Terms (TGMI), the Thesaurus for Graphic Materials II: Genre and Physical Characteristics Terms (TGM II). The Anglo-American Cataloguing Rules are used to describe scripts, manuscripts, partial notes and such. Some tracking information about other documentation is recorded using the Turabian Style Guide and The Chicago Manual of Style.
CS09(04), WGBH Boston
Current: In-house descriptive standards combined with modified Library of Congress Subject Headings. DAM: The above plus Dublin Core and PBCore (i.e., Public Broadcasting Core) compliant.

CS10, *Danube Exodus*

The interim report states that neither standards nor schemas are being used consistently in the environments studied. Forgács does capture metadata in the course of his work, but it is a system largely based on individual need, as informed by standard professional filmmaking practice. However, to date it is uncertain to the extent to which any metadata schema is currently used within the institution.

CS13, *Obsessed Again...*

There are no descriptive or other metadata schemas or standards currently being used.

CS15, *Waking Dream*

No descriptive or metadata standards are currently being used. There is no recordkeeping system being used.

23. What is the source of these descriptive or other metadata schema or standards (institutional conventions, professional body, international standard, individual practice, etc.?)

CS01, Arbo Cyber, théâtre (?)

Arbo does not use any descriptive or metadata standards. The report states that the "grey-zones" list information; thus, are not standardized.

CS02, Performance Artist Stelarc

The final report states that it is likely individual practice by Stelarc and his Web master that are the sources for any descriptive standards.

CS03, HorizonZero/ZeroHorizon Online Magazine & Media Database

The CanCore standard is derived from the Dublin Core metadata set, and is based on and fully compatible with the IEEE Learning Object Metadata standard and the IMS Learning Resource Meta-data specification. Other metadata sets are the result of individual practice.

CS09(01), Altair4 di Roma

The final report states that only material that is archived are then governed by international standards.

CS09(02), National Film Board

The NFB participates in international standards making bodies and is in some instances responsible for either assisting in developing these or in adapting them to the Canadian scene. These standards are, however, technical rather than descriptive.

CS09(03), Commercial Film Studio

Institutional convention governs practice during the workflow stage for any particular production. A snapshot of the entire directory structure for each production is kept, but users trying to access materials from even recent productions have been unsuccessful because of hardware and software changes that occurred in the meantime. Material that is archived is done so using the tools listed in the answer to Question 22, so professional bodies and international standards govern these activities.

CS09(04), WGBH Boston

Current: In-house data entry personnel with professional archives and library training, Library of Congress published and on-line sources. DAM: The above plus Dublin Core and PBCore (i.e., Public Broadcasting Core) reference resources.

CS10, *Danube Exodus*

The interim report states that this is not applicable.

CS13, Obsessed Again...

No such schema or standards are employed.

CS15, *Waking Dream*

Not applicable.

Focus 2. Scientific Activities

General information regarding metadata

CS06, Cybercartographic Atlas of Antarctica

The final report states that metadata in the field of geomatics are critical to business processes. The Cybercartographic Atlas of Antarctica acquires data from a number of organization and these data sets are accompanied by metadata (see Appendix K of the final report for details). The Atlas itself adheres to the ISO19115 geographic metadata standard for each module that has been entered into the MADRAS Registry developed at UCLA. Digital multimedia information objects (e.g., video clips, photos, audio, webcams, etc.) are also fully referenced and include metadata embedded into the object and/or accompanying the object and/or referenced as a caption and acknowledged in the bibliography of each content module. CS06 includes metadata-specific documents as follows:

- Excerpt - *Elements of geospatial data quality, March 8, 2002*
- *Multimedia Metadata Discussion Document, December 2003*
- Appendix P, List of Standards Adhered to on the Project

CS08, Mars Global Surveyor Data Records (NASA)

The PDS (Planetary Data System) uses self-describing data files as a preservation strategy. The labels of self-describing files describe the file format of attached data as well as the context in which the data were created. The PDS is referred to as an “active archive,” whereas the National Space Science Data Center’s (NSSDC’s) repository is referred to as a “deep archive.” The PDS is the entrance for Planetary Science data into the NSSDC archives for long-term preservation.

CS14, Archaeological Records in a GIS

The final report states that process for creating and maintaining the digital entities is ad hoc, even though GIS dynamically links geospatial data and descriptive attribute data from a wide variety of sources, and thus is a spatially referenced data set with specific metadata.

CS19, Preservation and Authentication of E-Engineering and Manufacturing Records

The main purpose of the engineering experiment examined by CS19 was to develop an open-source preservation format for digital computer-aided design (CAD) records of solid models used in high-tolerance manufacturing of complex assemblies. The experiment used Web Ontology Language (OWL), a W3C specification that extends XML to allow representation of semantics within metadata schemas, to persist the geometry, topology and functional characteristics of CAD model objects. The semantic format enabled automated querying of the digital entity’s meaning, expressed in its metadata in order to assess its authenticity. The CAD model objects were developed using proprietary reasoning programs and instantiated in accordance with ISO 10303, Standard for the Exchange of Product Model Data (STEP), AP 203 and part 21 EXPRESS. STEP is ISO’s metadata standard for the representation and electronic exchange of industrial product data between computer-based product life-cycle systems. AP 203 specifies the complete boundary representation of a solid model and EXPRESS defines its elements and attributes using an object-oriented approach (see 4a, below). Metadata elements were stored in the metadata catalogue management system (MCAT) of the ISO 14721, Open Archival Information System-compliant pilot preservation system managed by CS19 partners the Electronic Records Archives (ERA) Program of the U.S. National Archives and Records Administration (NARA), the University of Maryland and the San Diego Supercomputer Center (SDSC). This preservation system also incorporated SDSC's Storage Resource Broker technology, a middleware application that uses grid and metadata technologies to transparently manage data. The intent of the experiment was to preserve not only the geometric specifications of the model but also its semantically encoded metadata, joined to make a "new logical preservation format" for archival purposes. By logical preservation format, the experiment partners in CS19 meant a format encompassing not only the fixed form and content of information representing the model, but also instructions encoded within its metadata in a way that reasoning engines of the future can conduct "proofs" against the object to authenticate it as fit to support the procedural action for which it was designed to be used.

CS26, MOST Satellite Mission

The final report states that the MOST researchers chose file formats based upon best practice; thus, resulting in metadata based upon the file format chosen.

Metadata information in the 23 questions:

4a. What are the key formal elements, attributes, and behaviour (if any) of the digital entities?

CS06, Cybercartographic Atlas of Antarctica

The information expressed is primarily cartographic, according to the functionality of each of the file types below.

<u>Text</u>

- HTML
- XML with XSL style sheets
- Feedback / comment forum or blog
- Databases
- PostgreSQL—open source
- PostGIS (e.g., polygons, etc.)—open source
- Excel spreadsheet (scientific numeric data—e.g., local databases)
- ESRI EOO (e.g., Antarctic Digital database)
- Flat binary (e.g., National Snow and Ice Data Center (NSIDC) at NASA)
- Graphics (e.g., remote sensing data, terrain models, Digital Elevation Models
- (DEM), Radar data, pictures, etc.)
- 2-dimensional—BMP
- 2-dimensional—GIF
- 2-dimensional—JPEG
- 2-dimensional—TIF
- 2-dimensional—PNG
- 2-dimensional—GEOTiff

- 3-dimensional—VRML and the viewer(s) required to access it (e.g., Cortona or other that works with Firefox and Mozilla browsers)

Sound

- OGGVorbis—open source
- WAV
- AIF
- AU
- Moving images
- Quicktime
- MPEG4
- Animation
- SVG (Scalable Vector Graphics)—open standard
- Flash
- Virtual reality fly-through
- VRML or video
- Games
- Online quizz
- Programming languages and technical specifications
- Javascript
- Java
- SVG
- DHTML
- XML (schema files)
- GML (Geographic Markup Language)
- VRML (Virtual Reality Modeling Language)
- Haptics (e.g., a vibrating mouse, shaking chair, force feedback devices)
- Feasible if creator wishes to do so
- Operating System, Middleware
- Linux Redhat Enterprise V4
- Apache Server—open source
- TomCat - Java—open source
- PROJ—open source
- GEOS—open source
- Geoserver—open source
- Deegree—open source
- Java SDK—open source
- XML Libraries—open source
- WFS
- WMS/WCS

For additional details about the digital entities in use, please consult the following report appendices:

- Appendices H & J: Hardware and software lists
- Appendix J: Mime Encoding of Project Software
- Appendix K: List of Data Sources for the CAA

- Appendix M: Atlas Framework, Model and File Types - Freiburg Paper and Presentation

Given the complexity of the CAA, it is not possible to list all the digital components or their individual specifications. Please refer to Figure 2 in the report for a diagram of the overall technical architecture of the CAA.

CS08, Mars Global Surveyor Data Records (NASA)

The PDS data nomenclature standards define the rules for constructing Data Element and Data Object names. A keyword (standard data element name) is an element of the Planetary Science Data Dictionary (PSDD) that defines a named property of an object. The keyword plus its value is an attribute. A label (product label) is a resource description stored in a file. If the label is in the same file as the resource, it is called an attached label. If it is in a separate file, it is called a detached label. Labels also describe the structure or format of the data. Object Description Language (ODL) is used to create labels (data descriptions) for data files and other objects such as software and documents. The PDS labels contain the key attributes of the digital objects. The behaviours of a digital object consist of the various operations that can be performed on the object. For instance, an image object is a regular array of sample values. Image objects are normally processed with special display tools to produce a visual representation of the sample values. This operation on the digital object to produce a visual representation is a behaviour of an image object. Other behaviours of these digital objects consist of the processing and analytic tools that can be used to create other objects, e.g., a tool to produce an image histogram from an image.

CS14, Archaeological Records in a GIS

The final report states that the core data set is represented in both text and numeric characters, while the outputs are textual and graphic in nature (map(s) alongside tabulated data). Furthermore, the process for creating and maintaining these entities is ad hoc.

CS19, Preservation and Authentication of E-Engineering and Manufacturing Records

There were five (5) digital entities in the CS19 engineering experiment. The first two entities listed below are produced during actual computer-aided design (CAD) and computer-aided manufacturing (CAM) processes of the original experiment partner. In the actual business process, these entities are stored with a TIFF rendition of designs as an archival aggregate in a product data management system. They were extended in CS19's engineering experiment by three additional digital entities. Each iteration of format in the experiment was chosen to either strengthen semantic expressiveness or to capture knowledge representation in a persistent, open source format:

1. The original entities (1) are created by product designers using proprietary Pro-Engineer CAD systems and are provided to colleagues charged with computer-aided manufacturing of high-tolerance, high-assurance objects used in complex assemblies. There is no formal definition of this format in the public domain as the file has a proprietary format.
2. Corporate business rules of the original experiment partner ensure that the proprietary CAD design record (1) is translated into (2) Standard for the Exchange of Product Model Data (STEP) AP203 format (ISO 10303). The formal element, attribute and behaviour definition of the objects in the STEP file are contained in ISO 10303 AP 203. The standard describes the formal representation of the Euler complete boundary representation definition of a solid model. The definition of the elements and attributes are described in an object-oriented representation language called

EXPRESS that is ISO 10303 Part 21. EXPRESS schemas are computer-processable and can be verified automatically for syntactical correctness and for the existence of appropriate links to other schemas. Instances of the defined entities form the actual exchanged data. Entity definitions include rules that can be checked at translation time to verify certain aspects of semantic validity of the transferred instances.

3. From there, the experiment took the logical form of this STEP record (2) and enhanced it into another logical form (3) that supported the delineation of additional geometric relationships and reasoning about part shape and action or process semantics using C++ based knowledge representation tools. The derived features and action semantics able to be represented by this format allow for their automated interrogation by reasoning programs, establishing semantic metadata to enable automated archival authentication of the digital solid model.
4. These entities (3) were then taken through a proprietary reasoning engine (Logistica) to complete rendition of a format (4) with all required attributes and metadata, including the formulation of logical predicates. Although the Logistica format is proprietary, it can be said that it contains a knowledge component and a procedural reasoning component.
5. The Logistica entity (4) was converted to Web Ontology Language (OWL) format (5), an open source, public domain XML specification of the World Wide Web Consortium (W3C) for persistent archiving purposes. The OWL form is in ASCII. The logical components of this form are defined mathematically by concepts of descriptive logic, and the syntax of this form is defined by the W3C in the specifications. OWL is a semantic XML format to represent machine interpretable content when the content needs to be processed by applications rather than just structured for presentation to humans. This requirement applies not only to the World Wide Web but to the digital holdings of any given domain within it, including records repositories. OWL can be used to explicitly represent the meaning of terms in vocabularies and the relationships between those terms; in other words, to operationalize an ontology. OWL has more facilities for expressing meaning and semantics than XML, RDF and RDF-S, and thus OWL goes beyond these languages in its ability to represent machine interpretable content.

CS26, MOST Satellite Mission

The key elements are mainly textual, but there are graphic elements as well.

4d. How are the digital entities identified (e.g., is there a [persistent] unique identifier)?

CS06, Cybercartographic Atlas of Antarctica

There are no unique and/or persistent identifiers, and there is no formal ID lookup system.

- The digital objects are identified by a unique combination of a file name and a location in the system
- Some objects are identified in databases, with location information included with other metadata (see Question 22 below).
- There are also some metadata embedded within some digital objects. The modules are associated with metadata. Within a module, metadata are available to reference any entity via a citation.

- Some of the maps will have embedded Geographic Markup Language (GML) to link to and describe related geo-referenced objects, such as images or sounds.
- A multimedia metadata schema is being developed. Some of the elements will be embedded within the information objects themselves and some will be linked to the object. This will become a part of the Authors' Toolkit, which includes a template of the XML schema that is completed by the content creators.

CS08, Mars Global Surveyor Data Records (NASA)

A product ID data element represents a permanent, unique identifier assigned to a data product by its producer. In the PDS, the value assigned to a product ID must be unique within its dataset. The PDS Standards Reference also specifies the rules for dataset and volume names and IDs. Each PDS dataset must have a unique data set name and unique data set ID, both formed from up to seven components. Within datasets, there are unique volume IDs. Within volumes, the file names are unique.

CS14, Archaeological Records in a GIS

Digital entities are identified through file naming conventions. Aggregations of files within certain folders can also create an associative identity of their own.

CS19, Preservation and Authentication of E-Engineering and Manufacturing Records

In the business activities of the originating experiment partner, digital entities (1) and (2), along with a TIFF version of a solid model design, are stored according to documented company policies in a proprietary product data management system (PDM). The PDM in use is a commercial records management application. This aggregate, termed a "bill of materials," is filed in the PDM according to a numbered schema corresponding with design/manufacturing procedures and there under by project number. Within digital entities (3), (4) and (5), the underlying format allows the assignment of unique identifiers at the file level depending upon business needs. This is especially true of files formatted according to the ISO 10303 STEP AP203 and part 21 EXPRESS metadata schemas, which among their functions support specification of the bond between components in complex mechanical assemblies. It also should be noted that within individual CAD files and the semantic extension formats the representation of each individual attribute or element also has persistent unique identifiers. However, the protocol of the engineering experiment did not require the unique identification of each digital entity, since there was only one instance of each of the five entities. Furthermore, CS19 is founded on the proposition (already operational in the Semantic Web) that simple enumeration of discrete identity and integrity metadata is inadequate to the demands for discovery and authentication facing the future of archives. The conception of intrinsic documentary form needs to go much further into recognizing the characteristic patterns (classes, relationships, constraints) that cohere among and between otherwise static identity attributes.

CS26, MOST Satellite Mission

Digital entities are uniquely identified by file names [managed by 1) primary target (star) and 2) date]. In addition to this, the metadata provide another set of unique identifiers. The report does not explain what these identifiers are.

18b. From what applications do the recordkeeping system(s) inherit or capture all digital entities and the related metadata (e.g., e-mail, tracking systems, workflow system, office system, databases, etc.)?

CS06, Cybercartographic Atlas of Antarctica

The final report states that CAA relies on the XML-tagged content modules for the creation of metadata. CAA content modules are developed by content creators in such a way that the linkages between the information objects, their functionality and associated metadata are described in an XML document (created within the specified XML project schema), where the markup language indicates what to display. Subversion maintains all code, and all versions of that code are tracked. Subversion is from Tigris.org—an open-source content versioning system (CVS) for use with the most popular operating systems. The Subversion database is backed up regularly.

CS08, Mars Global Surveyor Data Records (NASA)

Instrument measurements are sent as data packets from spacecraft through the Deep Space Network to computers at the Mission Ground Station at JPL (Jet Propulsion Laboratory). Computer workstations of the various project institutions are connected via NASCOM and Ethernet links to a project database (PDB) at JPL. The workstations are used to create standard data products, documentation and index tables. These are packaged into archive volumes and sent to the Science Data Validation Team (SDVT) for validation. The SDVT transfers the archive volumes to the PDS where there is additional validation.

CS14, Archaeological Records in a GIS

The final report states that there is no recordkeeping system external from the applications; therefore, no formal capture activity. There are numerous capture activities within the GIS. Other than other elements of the Microsoft Office Suite, there are no collective capture tools for the information within the GIS. Groups of data are captured temporarily within the GIS application, ArcView while analysis is being conducted, but then is exported to its appropriate areas outside of the GIS application, either from Microsoft Excel or Access files.

CS19, Preservation and Authentication of E-Engineering and Manufacturing Records

In the business activities of the originating experiment partner the digital entities created in the CAD system are captured in the corporate PDM, which is a commercial records management application system (cf. 4d, above). The expression of the experiment digital entities into the final logical preservation format was a process of derivation and extension from both proprietary and open source systems, as detailed in 4a, above. Within the protocol of the CS19 engineering experiment, the digital entities and related metadata were captured by SDSC's Storage Resource Broker and NARA-ERA's Metadata Catalog Management System.

CS26, MOST Satellite Mission

The final report states that there is no formal capture system in place, beyond the tools within Microsoft Windows.

18d. Does the recordkeeping system provide ready access to all relevant digital entities and related metadata?

CS06, Cybercartographic Atlas of Antarctica

The final report states that a "multimedia metadata schema is being developed where some of the elements will be embedded within the information objects themselves and some will be linked to the object and these will become part of the Authors' Toolkit, which includes a template of the XML schema which is completed by the content creators." The ISO19115 metadata standard will be adhered to at the module level.

CS08, Mars Global Surveyor Data Records (NASA)

Presumably, but this is not clarified in the final report.

CS14, Archaeological Records in a GIS

No. As mentioned earlier, the recordkeeping environment is a dispersed and does not provide organized access. The creator is the intermediary between the files when access is needed, especially because the majority of the files are in the file directory or on the hard drive of the creator.

CS19, Preservation and Authentication of E-Engineering and Manufacturing Records

In the business context of the originating experiment partner the PDM system allows ready access to all digital entities and related metadata. Access is accomplished through standard queries invoked by menu picks by such attributes as procedure number, job, creator, design-change number, design release version number, etc. For the CS19 engineering experiment the SRB and MCAT systems provide a variety of means to access digital entities and any combination of metadata. In addition, the experiment protocol called for the logical querying of the semantic metadata of formats (3), (4) and (5), to authenticate the digital entity's identity, integrity and suitability for the manufacturing process for which it was designed.

CS26, MOST Satellite Mission

The final report reveals that it is possible to access all digital entities via Windows Explorer but does not actually mention how access is provided to the metadata prescribed by the MOST researchers.

18e. Does the recordkeeping system document all actions/transactions that take place in the system re: the digital entities? If so, what are the metadata captured?

CS06, Cybercartographic Atlas of Antarctica

Although this question is not directly applicable to CS06, answers to the following of the 23 research questions provided in the report do touch on this issue.

Question 8: Any digital object that forms part of the CAA must be described by the creator, using metadata standards adopted or developed by the project. See Question 20 in the report and Appendix P, which includes the project's metadata standards. Retrievability of, and access to, the digital objects are based on a number of adopted OGC interoperability specifications (see Appendices P and N in the report).

Question 10: Data are acquired from authoritative sources and are peer-reviewed (e.g., British Antarctic Survey, Scientific Committee on Antarctic Research, scientific and academic journals and books). Each is assessed against the Elements of Spatial Data Quality, which include:

- Lineage
- Positional accuracy
- Attribute/thematic accuracy
- Completeness
- Logical consistency

- Semantic accuracy
- Temporal information

See Appendices T and K in the report for the list of data sources.

Authenticity in geography is captured in standard metadata as data lineage. Quality measures are dependent on the type of data and their function (e.g., the acceptable margin of error for the precise location and size of a particular ice flow to inform tourist ships is smaller than fish counts to inform fisheries and ecological modeling). In addition, each scientific domain is governed by its particular data quality standards, measures and assurances and these are included in the metadata. Appendix P in the report includes a list of such standards.

Question 13: Changes to the code are captured in Subversion, a source repository system used by the project. Subversion maintains all code, and all versions of that code are tracked. Subversion is from Tigris.org—an open source content versioning system (CVS) for use with the most popular operating system. The Subversion database is backed up regularly. Other digital objects that form part of the CAA are not captured by Subversion.

The Authors' Toolkit will eventually allow changes to associated metadata to be tracked as well. Also see:

- Excerpt—*Elements of geospatial data quality, March 8, 2002*
- *Multimedia Metadata Discussion Document*, December 2003

CS08, Mars Global Surveyor Data Records (NASA)

The PDS logs accesses to restricted areas of the system. User ID, date, time and operation are logged.

CS14, Archaeological Records in a GIS

The report explicitly states that there is no audit trail. The GIS Specialist is in the process of creating metadata relating to the source of the data, including the original author, date or recording, etc.

CS19, Preservation and Authentication of E-Engineering and Manufacturing Records

The PDM system used by the originating research partner in actual business processes captures actions, events and changes to the digital entities (1), (2) and the bill of materials aggregate. Metadata is typically name of creator, release version numbers, date of release, etc. The SRB and MCAT systems captured all changes to the representation of the CAD solid model as it migrated through the semantic format extensions (3), (4) and (5), including the formulation of metadata that support querying by automated reasoning programs.

CS26, MOST Satellite Mission

The final report states that there is no audit trail.

22. What descriptive or other metadata schema or standard are currently being used in the creation, maintenance, use and preservation of the recordkeeping system or environment being studied?

CS06, Cybercartographic Atlas of Antarctica

The final report states that the CAA has solid metadata practices in place; these metadata practices include the following: FGDC and/or British Antarctic Survey DIF (Directory Interchange Format), OGC interoperability specifications and the International Standards

Organization 19115 Geomatics Standards. The report also indicates that the ISO 19115 metadata standard for digital mapping data has been explored (see *Multimedia Metadata Discussion Document, December 2003*).

CS08, Mars Global Surveyor Data Records (NASA)

The final report states that the *Planetary Science Data Dictionary* (PSDD) is used in the creation, maintenance, use and preservation of the PDS. The PSDD contains definitions of the standard data element names and objects.

CS14, Archaeological Records in a GIS

The final report states that they are interested in using ArcCatalogue, a metadata tool that is in the new version of ArcView. Their main goal relating to metadata capture surrounds source information relating to CC Database data. The metadata would indicate from what source (publication, repository, Web site, database, etc.) the data was retrieved. In addition, time tagging of georeferenced information is part of the documentation of the processes of creating online digital maps, models and georeferenced visualizations.

CS19, Preservation and Authentication of E-Engineering and Manufacturing Records

The final report states that the first digital entity (1), produced during actual computer-aided design (CAD) and computer-aided manufacturing (CAM) processes of the original experiment partner, originates in a proprietary software tool, thus the precise metadata schema is unavailable. However, the tool produces models in conformance with the ANSI Y-14.5 tolerance standard and provide export files (2) compliant with ISO 10303 Standard for the Exchange of Product Model Data (STEP), AP 203 and part 21 EXPRESS. Corporate metadata standards and procedures govern the filing of these two digital entities with a third TIFF export of the model view into a commercial Product Data Management System. The formats of CS19's digital entities (3) and (4) included the formulation of additional semantic metadata by in-house computer scientists expert in knowledge representation systems that supported the delineation of additional geometric relationships of the CAD solid model and reasoning about part shape and action or process semantics. Although some of the metadata supporting action semantics was lost in the translation to digital entity (5), OWL XML, it was able to persist and authenticate precise specifications about part shapes and relationships, including the classes, predicates and constraint rules that govern the identity and behavior of the CAD solid models.

CS26, MOST Satellite Mission

The metadata schema that is used was created by the MOST researchers and is specific for the data/files that are created in the MOST project. The metadata refer to information such as orbital parameters, observational parameters, telemetry information and target image information. The report notes that some of the metadata/descriptive fields in the FITS files are mandatory, due to the file format. In general, no metadata standards are used; the MOST researchers have created their own scheme of important descriptive fields.

23. What is the source of these descriptive or other metadata schema or standards (institutional conventions, professional body, international standard, individual practice, etc.?)

CS06, Cybercartographic Atlas of Antarctica

The source of metadata comes directly from professional bodies, institutional conventions, as well as international standards. The Atlas adheres to ISO 19115 at the modular level and additional research is ongoing regarding metadata at the granular level.

See response to question 22, above.

- International Standards Organization (ISO)
- Open Geospatial Consortium (OGC)
- Scientific Committee on Antarctic Research (SCAR)
- Geomatics and Cartographic Research Centre (GCRC)
- DIF Format (see http://gcmd.gsfc.nasa.gov/User/difguide/difman.html for details)

The CAA project itself: Y. Zhou, MA thesis on this topic entitled "Profiling and Visualizing Metadata for Multimedia Information in a Geospatial Portal."

CS08, Mars Global Surveyor Data Records (NASA)

The *Planetary Science Data Dictionary* is a NASA institutional standard for Planetary Science Metadata. The PDS procedures for assigning standardized names to data elements follow closely the NBS Guide on Data Entity Naming Conventions.

CS14, Archaeological Records in a GIS

Within ArcCatalogue, the user could create, manage and edit metadata based on the Federal Geographic Data Committee (FGDC) Content Standards for Digital Geospatial Metadata or the ISO 19115 Metadata Standard. These metadata would be stored in XML.

CS19, Preservation and Authentication of E-Engineering and Manufacturing Records

ANSI, ISO, W3C, corporate business rules

CS26, MOST Satellite Mission

The metadata that are used for the various files are based on experience and best practice in the astronomical community and on the foreseeable use of the records in the future. There is an internal MOST document that describes the descriptive fields of the FITS files.

Focus 3. Governmental Activities

General information regarding metadata

CS05, Archives of Ontario Web Exhibits

The final report states that the Ontario government has developed a standard look and feel to which all government Web content must adhere. These are standards created or adopted within the Ontario Public Service. One such standard is the metadata, which refers to title, keyword and description and classification metatags.

CS12, Antarctic Treaty Searchable Database

The final report states this "is an automated technology that objectively integrates digital-record entities without markup, metadata or databases." However, the report further states that "unlike subjective content descriptions in metadata or controlled vocabularies, DIGIN® comprehensively searches both the contents of the granules and their categorical tags to objectively identify those granules that match the search queries. DIGIN® is interoperable with metadata, mark-up and databases."

CS17, New York State DMV On-line Services System

The final report states that the DMV provides a highly interactive online system featuring a complex set of interwoven electronic activities, which collects information about the user via cookies, Web protocols and transactional metadata. A third party digital signature company, VeriSign, is used to make transactions legally binding.

CS18, Alsace-Moselle's Land Registry

The final report states that there are no descriptive or other metadata schemas or standards; however, within the relational database the data are linked together through queries. It has been explained that there is a metadata schema that will be completed for the second phase of the project.

CS20, Revenue On-Line Service (ROS)

The final report states that data mining of ROS-created data is used to audit tax details, improve efficiencies, increase customer service and enable fraud detection.

CS21, Electronic Filing System, Supreme Court of Singapore

The law firm is expected to enter information on their cases through a prescribed documentary template in EFS. Some of the metadata elements are fixed as there is a pull-down menu for law firms to select. Some of the metadata elements the file has to enter include the firm's file reference number and party details, which include the party type, name and address of parties and name of solicitor. EFS captures both the metadata of the record and the actual record itself; the court must check both the metadata and the record.

CS24, City of Vancouver GIS (VanMap)

The interim report states that metadata is not a means of tracking how the information is used, but it does reveal what information is being used and when; this is conducted through the generation of statistical reports.

CS25, Legacoop of Bologna Web Site

The final report explains that all paper mail (what is sent to the organization and what is sent directly to the single functionaries) is registered. And, that the electronic mail sent to the organization official email-system is registered when it is determined that the message is of a certain importance. The registry system uses an automated application to register the records. This application provides a profile of the registered incoming and outgoing documents, including the following: classification code, recipients, object, date and type of document.

Metadata information in the 23 questions:

4a. What are the key formal elements, attributes, and behaviour (if any) of the digital entities?

CS05, Archives of Ontario Web Exhibits

Elements and attributes that are considered integral to the validity and completeness to the Web exhibits Elements were determined based on how the exhibits are normally accessed.

Key intrinsic elements include:

- navigation links from the institutional home page to a listing (with or without a précis of the exhibit);
- exhibit content, normally comprised of Web pages containing text, images, and occasionally with sound or video files;
- government visual identity signs, especially the provincial and city logos and the institutional;
    - Provided by a central body for all Ontario Web sites are:
        - Standard disclaimers
        - Instructions for accessing and installing plug-ins
        - Copyright statements
        - Privacy statements
        - Graphics (.gif format) provided for every ministry name

- Graphics for mandatory toolbars are provided
- Ontario logo, mandatory for every government Web page, and footer graphics are provided

The last three are compliant with the W3C's WAI (Web site Accessibility Initiative) requirements, and all text is provided in English and French.

Key extrinsic elements include:

- A corporate standard Web page template
- The cascading style sheet created for the Web site as a whole
- The institutional Web site (contains other exhibits, links to databases, external links, etc.)
- The corporate Web environment (contains links to all government Web sites, news releases, etc.)
- HyperText Markup Language, specification version 4.01
- Navigation bars required at the top/bottom/side of each Web page
- A "feedback form" that utilizes Common Gateway Interface (CGI) script to interface with an email application

Behaviour of the rendering platform takes place on two levels:

13. the feedback form is a CGI program executed in real-time; and
14. the way the user's browser interacts with the HTML coding of the exhibits.

CS12, Antarctic Treaty Searchable Database

Indeterminate from answer provided.

CS17, New York State DMV On-line Services System

The records are live records and have the ability to change over time. They can be placed into a status where they are no longer alterable, as when a driver dies or a vehicle is junked. The official records contain an official crest or logo.

CS18, Alsace-Moselle's Land Registry

The ordonnances take the form of XML files, containing tagged information relative to landowners, land parcels and rights/obligations relative to the property. Associated with the ordonnance is a digital signature of the judge authoring the ordonnance. The structure of the ordonnance is defined through a DTD. The scanned images of the register take the form of TIFF files.

CS20, Revenue On-Line Service (ROS)

There are some elements and attributes common to the three "classes" of records [digital certificates/signatures, tax forms and debit instructional forms] for presentation, Revenue logo, font and style, certification practice statement, privacy policy, terms and conditions, copyright statements, contact details and standard Web page templates.

CS21, Electronic Filing System, Supreme Court of Singapore

The EFS is composed of standardized HTML style sheets, XML files, Visual Basic, PDF and graphic files for the EFS logo.

CS24, City of Vancouver GIS (VanMap)

The data sheets describe VanMap's data layers, features and functionalities; each layer typically contains, layer name, group name, scale, data currency status, responsible department and definition.

CS25, Legacoop of Bologna Web Site

All the entities have at least a title and a body text and a date. Each element is numbered sequentially according to the chronological order.

4d. How are the digital entities identified (e.g., is there a [persistent] unique identifier)?

CS05, Archives of Ontario Web Exhibits

Each Web exhibit is identified by its title and a URL, which has been assigned within the institution's Web domain. When viewing the source coding for each Web page within each exhibit, each page is also titled.

CS12, Antarctic Treaty Searchable Database

Each of the information granules or digital-record entities in the current *Database* contains unique provenance information in a categorical header tag as well as in the title. Unlike metadata, which are stored in repositories separately from the digital entities, the unique identifiers are part of each granule in the *Database*. Thus, with the categorical header tags, there is never a risk for decoupling the unique identifiers and the digital entities.

CS17, New York State DMV On-line Services System

There is a unique identifier connected to each transaction. The transaction and its identifier are stored with the core record, as a result of the transaction. Different sets of identifiers exist for each of the three file types: license, registration, and title.

CS18, Alsace-Moselle's Land Registry

Every inscription in the database is numbered with a persistent, unique identifier and dated. Ordonnances are also numbered and dated. Each scanned image of the registers is numbered according to the system already in place for numbering individual pages of the registers.

CS20, Revenue On-Line Service (ROS)

There is no need for specialized codes and keys beyond those normally used by the Revenue.

CS21, Electronic Filing System, Supreme Court of Singapore

The case number is a unique identifier, which is automatic generated number assigned by the courts.

CS24, City of Vancouver GIS (VanMap)

The HTML and CML pages and embedded GIF images are identified by unique URLs. The data fields, layers and groups are also identified by field names, layer names and group names.

CS25, Legacoop of Bologna Web Site

A primary key in the form of a progressive number (managed as a key field in the database) is the main identification attribute.

4e. In the organization of the digital entities, what kind of aggregation levels exist, if any?

CS05, Archives of Ontario Web Exhibits

The Web exhibits and the Web pages reflect aggregations of text, images and other components of the exhibit which are conceptually linked. The institutional Web sites and the Web exhibits are grouped together for the navigational convenience of the user.

CS12, Antarctic Treaty Searchable Database

The aggregation levels among digital entities are based on the inherent parent-child relationships within the policy documents. In general, the aggregation levels or hierarchy levels reflect the granularity of a digital collection. This collection granularity is represented specifically by:

- Antarctic Treaty Searchable Database > Year > Major Document or Antarctic Treaty Consultative Meeting > "measures"

Dynamic aggregation of digital-record entities with DIGIN® facilitates the discovery of relationships within and between the digital-record series.

CS17, New York State DMV On-line Services System

The DMV does not use directories or subdirectories, but keeps everything in tables and databases. The individual transactions in the audit trail are organized by date and time, category and current status.

CS18, Alsace-Moselle's Land Registry

The database aggregates the data according to the main categories: parcels, persons, rights and obligations. The presentation of data is organized in the same way as the paper register; that is, a single *feuillet* contains information relative to all the properties of a person within a given administrative territory (usually, a commune, or part thereof).

CS20, Revenue On-Line Service (ROS)

All tax records and debit instructions are saved chronologically and are viewable within the Revenue Customer Information Service. They can be sorted and viewed depending upon the field type selected. Regarding digital certificates and signatures: Metadata surrounding the older Digital Certificates, in addition to the security wrapper, are maintained with ROS. Revenue has a separate Archiving Policy for Certificates, but this is considered beyond the remit of ROS.

CS21, Electronic Filing System, Supreme Court of Singapore

The main case files are divided onto various sub-folders based on the type and nature of records filed, such as affidavit, draft order, minute sheet and summon in chambers.

CS24, City of Vancouver GIS (VanMap)

HTML and related pages are grouped into folders for storage, identification and retrieval purposes. The data are organized into layers, with each layer including a single data source or multiple data sources.

CS25, Legacoop of Bologna Web Site

The entities are aggregated according to the main logical categories of the Web site (documents of the association, news from the cooperation world, CVs and announcements, other services related to the Bologna business area).

18b. From what applications do the recordkeeping system(s) inherit or capture all digital entities and the related metadata (e.g., e-mail, tracking systems, workflow system, office system, databases, etc.)?

CS05, Archives of Ontario Web Exhibits

The exhibits are created using Dreamweaver and PageMaker software applications.
Metadata captured would normally be what are automatically captured by the default settings of those applications. None of the interviewees commented that they used the document properties function to add any specific metadata. Metadata captured would normally be what are automatically captured by the default settings of the applications used to create supporting documentation, such as Microsoft Word.

CS12, Antarctic Treaty Searchable Database

According to the final report, this question is irrelevant since, after the initial implementation of the *Database* in 1999, the only captured files are the entire Antarctic Treaty Consultative Meeting (ATCM) Final Reports without metadata that have been published on the ATCM Web sites of the host nations. The new *"measures"* that have been adopted by the Antarctic

Treaty Consultative Parties are then extracted and added to the *Database* with header tags that define their unique location in the overall collection.

CS17, New York State DMV On-line Services System

The system that the DMV uses captures IP addresses, system dates and session cookies. The cookies are used only to maintain the session state; they are not stored on the hard drive of the patron.

CS18, Alsace-Moselle's Land Registry

Both ordinances and inscriptions are captured through custom applications. The scanned images of the register were captured once at the onset of the computerization process.

CS20, Revenue On-Line Service (ROS)

Databases and other systems – ITP is held on an Ingress II mainframe back-end system.
ITP – Integrated Taxation Processing [System]

CS21, Electronic Filing System, Supreme Court of Singapore

The EFS captures digital entities from an oracle database, Filenet (document management system), jukbox (for CDs) and visual basic software. See answer to question 5a in final report for a detailed list of the application systems.

CS24, City of Vancouver GIS (VanMap)

DOMINO, PRISM, License and other systems produce the data.

CS25, Legacoop of Bologna Web Site

The report states that the recordkeeping system does not have any relationships with the Web site system.

18d. Does the recordkeeping system provide ready access to all relevant digital entities and related metadata?

CS05, Archives of Ontario Web Exhibits

The report states that related metadata is not readily accessible, even if it has been captured. This is due to the absence of a recordkeeping system and lack of consistent recordkeeping processes around the provision of access to Web exhibits within the two institutions.

CS12, Antarctic Treaty Searchable Database

The final report says yes, through providing comprehensive integrated access to the digital-record entities. And, that the Antarctic Treaty Searchable Databases does not require metadata.

CS17, New York State DMV On-line Services System

If the mainframe system is equivalent to the recordkeeping system, then the answer to this question is yes. Although customers and third party users have access to only a small portion of the digital entities, the system provides DMV personnel with access to all aspects of the digital entities.

CS18, Alsace-Moselle's Land Registry

Yes, as relevant for each category of user.

CS20, Revenue On-Line Service (ROS)

Yes, to both Revenue employees and ROS users. Not all users will view metadata.

CS21, Electronic Filing System, Supreme Court of Singapore

Yes, as relevant for each category of user. Authorized court users can view both the record profile and the PDF record. Specified group of users can, based on their job competency, view certain categories of audit logs.

CS24, City of Vancouver GIS (VanMap)

Metadata in the form of data sheets is also readily available. This study has not yet investigated the types of metadata that may be generated automatically by the various technological processes used to create VanMap.

CS25, Legacoop of Bologna Web Site

Not applicable.

18e. Does the recordkeeping system document all actions/transactions that take place in the system re: the digital entities? If so, what are the metadata captured?

CS05, Archives of Ontario Web Exhibits

The Web logging software documents aspects of all interactions with the institution's Web site. The report presents 21 reports generated by *Analog* based on the data it gathers. Please refer to page 46 of the report for this list.

CS12, Antarctic Treaty Searchable Database

The report states metadata are not captured. However, the report also states that all queries of the Web site version of the *Database* are automatically logged.

CS17, New York State DMV On-line Services System

The recordkeeping system at the DMV tracks all changes to records in the mainframe system through audit trails and user logs.

CS18, Alsace-Moselle's Land Registry

The report reveals that the system includes extensive login capabilities for recording all actions and transactions taking place in the system. Logs may be used for two distinct purposes.

CS20, Revenue On-Line Service (ROS)

The report states that all changes are noted and logged with time/date stamp and name of Revenue employee making change.

CS21, Electronic Filing System, Supreme Court of Singapore

Yes, all actions and transactions are documented in various audit logs, including:

- Transaction Log: Records user ID of user who activates the change, function name, date./time of the change, data items before and after the change
- Financial Audit Log: Records user ID, function name, date/time of the action, case number/document, control number/unique reference number, amount of fees before the change, amount of fees after the change, remarks, approval for exemption/waiver of court filing fees, approval for request of waiver of hearing fees, and approval for refund of hearing fees.
- Violation Log: Records user ID of user who attempts to access functions he or she is not granted access to, unsuccessful log in attempts, function name, date/time of the action, and brief description of the nature of the violation.

CS24, City of Vancouver GIS (VanMap)

The interim report states that the use of the data can be tracked by unique client IDs randomly generated when users download the MapGuide ActiveX Viewer to their workstations. For example, whenever a user accesses VanMap and issues a request for data the transaction results in a log file record containing his or her ID, the date and time of access and strings of numbers representing specific data layers used.

CS25, Legacoop of Bologna Web Site

Not applicable.

22. What descriptive or other metadata schema or standard are currently being used in the creation, maintenance, use and preservation of the recordkeeping system or environment being studied?

CS05, Archives of Ontario Web Exhibits

The Web site coordinator was unfamiliar with metatags, and initially ignored metadata standards. Metatags are only applied to "key pages"; therefore, is nothing that distinguishes an exhibit page from any other page on the Web site. The report states that only comprehensive source of metadata governing the entirety of an exhibit appears to be the "definition document" created for *The War of 1812* exhibit. This document includes the title, reference code, image number (where applicable), location information and a summary of the document/image.

CS12, Antarctic Treaty Searchable Database

Descriptive metadata, as conventionally applied with templates and attributes that reside in repositories, are not used to implement the Antarctic Treaty Searchable Database. However, there is the use of header tags that describe the parent-child provenance. Also, conventional metadata regarding the portal for the Antarctic Treaty Searchable Database are being added to the National Science Digital Library and Digital Library for Earth System Education. The metadata format for these submissions is a modified Dublin-Core metadata with additional fields for the education audiences that are being addressed by these digital libraries.

CS17, New York State DMV On-line Services System

Although data layouts and schema are used, the DMV respondents indicated that they did not feel comfortable revealing specifics about such information to the InterPARES research team.

CS18, Alsace-Moselle's Land Registry

No descriptive nor metadata schema is currently being used.

CS20, Revenue On-Line Service (ROS)

Twenty-two schemas for the tax forms have been made available in XML DTDs for inclusion in the third party compatible software; view www.ros.ie/downloads.html and Appendix IV. All schemas include a DTD and element definitions and explanations. Although an Irish Public Service Metadata standard exists, it is not used with ROS.

CS21, Electronic Filing System, Supreme Court of Singapore

The schemas for the documentary templates of the EFS are based on the workflow and juridical requirements of the court.

CS24, City of Vancouver GIS (VanMap)

The metadata applied by the VanMap developers include layer name; group name; scale at which the data become available; data currency status; responsible department, branch or division; and definition. Not all of these metadata are applied to all of the data layers. Metadata generated automatically upon creation of the data have not yet been investigated.

CS25, Legacoop of Bologna Web Site

Basic metadata related to the registry system are required in the recordkeeping system, but not exported to the Web site management, which handles only a numbering system for each digital entity and a date.

23. What is the source of these descriptive or other metadata schema or standards (institutional conventions, professional body, international standard, individual practice, etc.?)

CS05, Archives of Ontario Web Exhibits

There is no identified source for the Government of Ontario Category Metadata rules. The City's Web coordinator stated that the metadata tags he uses do not conform to any standards.

CS12, Antarctic Treaty Searchable Database

The final report states that conventional metadata are unnecessary with the DIGIN® technologies, which can interoperate with or without metadata to integrate "structured" as well as "unstructured" information. The sources of the descriptive schema are the persistent digital-record entities themselves.

CS17, New York State DMV On-line Services System

The source of these standards was not mentioned or discussed during the case study interview.

CS18, Alsace-Moselle's Land Registry

No descriptive nor metadata schema is currently being used.

CS20, Revenue On-Line Service (ROS)

Institutional practice. Form design and structure is based on existing paper-based forms. Field selection and management is based on requirements and format of ITP applications and data flow to this and other back-end systems. The XML schemas may include other descriptive standards such as ISO Year Standard.

CS21, Electronic Filing System, Supreme Court of Singapore

Institutional practice. The metadata used in the documentary template are based on common data elements associated with the court records that have to be converted into PDF.

CS24, City of Vancouver GIS (VanMap)

Metadata are based on what the VanMap Team thinks will be useful information for the end user.

CS25, Legacoop of Bologna Web Site

The metadata included are strictly related to the professional standard followed for building the Web site (SQL for the database and HTML for the Web site pages).

APPENDIX 19

A Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records

by

Luciana Duranti, The University of British Columbia
Jim Suderman, City of Toronto Archives
Malcolm Todd, National Archives of the United Kingdom

A Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records[1]

Introduction

The InterPARES research projects have examined the creation, maintenance and preservation of digital records. A major finding of the research is that, to preserve trustworthy digital records (i.e., records that can be demonstrated to be reliable, accurate and authentic), records creators must create them in such a way that it is possible to maintain and preserve them. This entails that a relationship between a records creator[2] and its designated preserver[3] must begin at the time the records are created.[4]

The InterPARES 1 research (1999-2001) was undertaken from the viewpoint of the preserver. Three central findings emerged from it: 1) there are several requirements that should be in place in any recordkeeping environment aiming to create reliable and accurate digital records and to maintain authentic records;[5] 2) it is not possible to preserve digital records but only the ability to reproduce them;[6] and 3) the preserver needs to be involved with the records from the beginning of their lifecycle to be able to assert that the copies that will be selected for permanent preservation are indeed authentic copies of the creator's records.

The InterPARES 2 research (2002-2006) took the records creator's perspective. The researchers carried out case studies of records creation and maintenance in the artistic, scientific and governmental sectors; they modeled the many functions that make up records creation and maintenance and records preservation according to both the lifecycle and the continuum models; they reviewed and compared legislation and government policies from a number of different countries and at different levels of government, from the national to the municipal; they analyzed many metadata initiatives and developed a tool to identify the strengths and weaknesses of existing metadata schemas in relation to questions of reliability, accuracy and authenticity; and,

[1] The term initially used in the InterPARES Project is "electronic records." In fact, the book resulting from InterPARES 1 is named *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* (Luciana Duranti, ed.; San Miniato, Archilab, 2005), and the formal title of InterPARES 2 carries that terminology forward. However, in the course of the research, the term "electronic record" began to be gradually replaced by the term "digital record," which has a less generic meaning, and by the end of the research cycle, the research team had developed separate definitions for the two terms and decided to use the latter as the one that better describes the object of InterPARES research. The definition for "electronic record" reads: "An analogue or digital record that is carried by an electrical conductor and requires the use of electronic equipment to be intelligible by a person." The definition for "digital record" is, effectively, a digitally-encoded object and the metadata necessary to order, structure or manifest the object's content and form, where "digital object" is taken to mean "a discrete aggregation of one or more bitstreams and the metadata about the properties of the object and, if applicable, methods of performing operations on the object." See the InterPARES 2 Terminology Database, available at http://www.interpares.org/ip2/ip2_terminology_db.cfm.

[2] Records creator is the physical or juridical person (i.e., a collection or succession of physical persons, such as an organization, a committee, or a position) who makes or receives and sets aside the records for action or reference. As such, the term includes all officers who work for a juridical person, such as records managers, records keepers and preservers.

[3] Records preserver is a generic term that refers more to the function than to the professional designation of the physical or juridical person in question. Thus, the preserver might be a unit in an organization, a stand-alone institution, an archivist or anyone else who has as primary responsibility the long-term preservation of records.

[4] Records are created when they are made or received and set aside or saved for action or reference.

[5] See Authenticity Task Force (2002). "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 204–219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf.

[6] See Kenneth Thibodeau et al., "Part Three – Trusting to Time: Preserving Authentic Records in the Long Term: Preservation Task Force Report," ibid, 99–116. Online reprint available at http://www.interpares.org/book/interpares_book_f_part3.pdf.

once again, they studied the concept of trustworthiness and its components, reliability, accuracy and authenticity and how it is understood, not just in the traditional legal and administrative environments, but in the arts, in the sciences and in the developing areas of e-government.

The case studies showed that record creation in the digital environment is almost never guided by considerations of preservation over the long term. As a result, the reliability, accuracy and authenticity of digital records can either not be established in the first place or not be demonstrated over periods of time relevant to the "business"[7] requirements for the records. These records cannot therefore support the creator's accountability requirements, nor can they be effectively relied upon either by the creator for reference or later action or by external users as sources. Furthermore, they cannot be understood within an historical context, thereby undermining the traditional role of preserving organizations such as public archival institutions.

The research undertaken in records and information-related legislation showed that no level of government in any country to date has taken a comprehensive view of the records lifecycle, and that, in some cases, legislation has established significant barriers to the effective preservation of digital records over the long term, most notably that regarding copyright.

It was the responsibility of the InterPARES 2 Policy Cross-domain research team (hereinafter "the Policy team") to determine whether it was possible to establish a framework of principles that could guide the creation of policies, strategies and standards, and that would be flexible enough to be useful in differing national environments, and consistent enough to be adopted in its entirety as a solid basis for any such document. In particular, such a framework had to balance different cultural, social and juridical perspectives on the issues of access to information, data privacy and intellectual property.

The findings of the InterPARES 1 research were confirmed by the research conducted by the InterPARES 2 Policy team, which further concluded that it is possible to develop such a framework of principles to support record creation, maintenance and preservation, regardless of jurisdiction. This document, in combination with other products of the Project, especially the Chain of Preservation (COP) model,[8] reflects this conclusion, while emphasizing the need to make explicit the nature of the relationship between records creators and preservers.

The Policy team developed two complementary sets of principles, one for records creators and one for records preservers, which are intended to support the establishment of the relationship between creators and preservers by demonstrating the nature of that relationship.[9] The principles for records creators are directed to the persons responsible for developing policies and strategies for the creation, maintenance and use of digital records within any kind of organization, and to national and international standards bodies. The principles for records preservers are directed to the persons responsible for developing policies and strategies for the long-term preservation of digital records within administrative units or institutions that have as their core mandate the preservation of the bodies of records created by persons, administrative units or organizations external to them, selected for permanent preservation under their jurisdiction for reasons of legal, administrative or historical accountability. They are therefore intended for administrative units (e.g., a bank, a city or a university archives) or institutions (e.g.,

---

[7] The term "business" is used in its most general sense, since the object of the InterPARES research includes works of art and scientific data as well as standard types of business records.
[8] The COP model is available in Appendix 14 and at http://www.interpares.org/ip2/ip2_models.cfm. A narrative discussion of the model is provided in the Modeling Cross-domain Task Force Report.
[9] The initial draft of the principles relied heavily on the contributions of three research assistants: Fiorella Foscarini, Emily O'Neill and Sherry Xie.

a community archives or a state archives) with effective knowledge of records and records preservation.

Structure of the Principles

The principles are similarly presented, with the principle statement followed by an explanatory narrative, sometimes with illustrative examples. The principles are more often phrased as recommendations ("should") rather than imperatives ("must"), because some of them might not be relevant to some records creators or preservers. Each principle statement is followed by an indication of the corresponding principle in the other set (C stands for Creator, P stands for Preserver; the number is the principle number in the C or the P set). The reason why the principle numbers do not correspond in the two sets (C1=P1) is that the principles are listed in each set in order of relative importance.

Principles for Records Creators

(C1) Digital objects must have a stable content and a fixed documentary form to be considered records and to be capable of being preserved over time. (P5)

The InterPARES Project has defined a record as "a document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference,"[10] adopting the traditional archival definition. This definition implies that, to be considered as a record, a digital object generated by the creator must first be a document; that is, must have stable content and fixed documentary form. Only digital objects possessing both are capable of serving the record's memorial function.

The concept of *stable content* is self-explanatory, as it simply refers to the fact that the data and the information in the record (i.e., the message the record is intended to convey) are unchanged and unchangeable. This implies that data or information cannot be overwritten, altered, deleted or added to. Thus, if one has a system that contains fluid, ever-changing data or information, one has no records in such a system until one decides to make one and to save it with its unalterable content.

The concept of *fixed form* is more complex. A digital object has a fixed form when its binary content is stored so that the message it conveys can be rendered with the same documentary presentation it had on the screen when first saved. Because the same documentary presentation of a record can be produced by a variety of digital formats or presentations,[11] fixed form does not imply that the bitstreams must remain intact over time. It is possible to change the way a record is contained in a computer file without changing the record; for example, if a digital object generated in '.doc' format is later saved in '.pdf' format, the way it manifests itself on the screen—its documentary presentation, or "documentary form"—has not changed, so one can say that the object has a fixed form.

One can also produce digital information that can take several different documentary forms. This means that the same content can be presented on the screen in several different ways, the various types of graphs available in spreadsheet software being one example. In this case, each presentation of such a digital object in the limited series of possibilities allowed by the system is to be considered as a different view of the same record having stable content and fixed form.

In addition, one has to consider the concept of "bounded variability," which refers to changes to the form and/or content of a digital record that are limited and controlled by fixed rules, so that the same query, request or interaction always generates the same result.[12] In such cases, variations in the record's form and content are either caused by technology, such as different operating systems or applications used to access the document, or by the intention of the author or writer of the document. Where content is concerned, the same query will always return the same subset, while, as mentioned, its presentation might vary within an allowed range, such as

[10] See InterPARES 2 Terminology Database, op. cit.

[11] Digital format is defined as "The byte-serialized encoding of a digital object that defines the syntactic and semantic rules for the mapping from an information model to a byte stream and the inverse mapping from that byte stream back to the original information model" (InterPARES 2 Terminology Database, op. cit.). In most contexts, digital format is used interchangeably with digital file-related concepts such as file format, file wrapper, file encoding, etc. However, there are some contexts, "such as the network transport of formatted content streams or consideration of content streams at a level of granularity finer than that of an entire file, where specific reference to "file" is inappropriate" (Stephen L. Abrams (2005), "Establishing a Global Digital Format Registry," *Library Trends* 54(1): 126. Available at http://muse.jhu.edu/demo/library_trends/v054/54.1abrams.pdf).

[12] See Duranti and Thibodeau, "The Concept of Record," op. cit.

image magnification. In consideration of the fact that what causes these variations also limits them, they are not considered to be violations of the requirements of stable content and fixed form.

Organizations should establish criteria for determining which digital objects need to be maintained as records and what methods should be employed to fix their form and content if they are fluid when generated. The criteria should be based on business needs but should respect as well the requirements of legal, administrative and historical accountability.

(C2) Record creation procedures should ensure that digital components of records can be separately maintained and reassembled over time. (P4)

Every digital record is composed of one or more digital components. A digital component is a digital object that is part of one or more digital records, including any metadata necessary to order, structure or manifest content, and that requires a given preservation action. For example, an e-mail that includes a picture and a digital signature will have at least four digital components (the header, the text, the picture and the digital signature). Reports with attachments in different formats will consist of more than one digital component, whereas a report with its attachments saved in one PDF file will consist of only one digital component. Although digital components are each stored separately, each digital component exists in a specific relationship to the other digital components that make up the record.

Preservation of digital records requires that all the digital components of a record be consistently identified, linked and stored in a way that they can be retrieved and reconstituted into a record having the same documentary presentation it manifested when last closed. Each digital component requires one or more specific methods for decoding the bitstream and for presenting it for use over time. The bitstream can be altered, as a result of conversion for example, as long as it continues to be able to fulfil its original role in the reproduction of the record. All digital components must be able to work together after they are altered; therefore, all changes need to be assessed by the creator for the effects they may have on the record.

Organizations should establish policies and procedures that stipulate the identification of digital components at the creation stage and that ensure they can be maintained, transmitted, reproduced, upgraded and reassembled over time.

(C3) Record creation and maintenance requirements should be formulated in terms of the purposes the records are to fulfil, rather than in terms of the available or chosen record-making or recordkeeping technologies. (P6)

Digital records rely, by definition, on computer technology and any instance of a record exists within a specific technological environment. For this reason, it may seem useful to establish record creation and maintenance requirements in terms of the technological characteristics of the records or the technological applications in which the records may reside. However, not only do technologies change, sometimes very frequently, but they are also governed by proprietary considerations established and modified at will by their developers. Both these factors can significantly affect the accessibility of records over time. For these reasons, references to specific technologies should not be included in records policies, strategies and standards governing the creation and maintenance of an organization’s records. Only the business requirements and obligations that the records are designed to support should be explicitly kept in consideration at such a high regulatory level. At the level of implementation, the characteristics of specific technologies should be taken into account to support the established business requirement and make possible its realization.

Technological solutions to record creation and maintenance are dynamic, meaning that they will evolve as the technology evolves. New technologies will enable new ways of creating records that meet an organization's business requirements. The rapid adoption of Web technologies to support business communication and transaction illustrates this. Specific activities for maintaining records will therefore require continuing adaptation to new situations drawing on expertise from a number of disciplines. To extend the example of the use of Web technologies, organizations creating and maintaining transactional records in a mainframe environment need to draw on knowledge of the new Web technologies from both connectivity (i.e., how to connect the mainframe to the Web) and security standpoints (i.e., how to protect the records from remote, Web-based attacks). As new technologies are used to create records, reference to new archival knowledge will continue to be required.

Technological solutions need to be specific to be effective. Although the general theory and methodology of digital preservation applies to all digital records, the maintenance solutions for different types of records require different methods. Therefore, they should be based on the specific juridical-administrative context in which the records are created and maintained, the mandate, mission or goals of their creator, the functions and activities in which the records participate and the technologies employed in their creation to ensure the best solutions are adopted for their maintenance.

Record policies that are expressed in terms of business requirements rather than technologies will need to be periodically updated as the organization's business requirements change, rather than as the technology changes. It is the role of a specific action plan to identify appropriate technological solutions for the maintenance of specific aggregations of records. The identified solutions must be monitored with regard to the possible need for modifying and updating. This requires the records creating body to be aware of new research developments in the archival and records management fields and to collaborate with interdisciplinary efforts to develop appropriate methods for the management of digital records.

(C4) Record creation and maintenance policies, strategies and standards should address the issues of record reliability, accuracy and authenticity expressly and separately. (P2)

In the management of digital records, reliability, accuracy and authenticity are three vital considerations for any organization that wishes to sustain its business competitiveness and to comply with legislative and regulatory requirements. These considerations should be directly and separately addressed in records policies and promulgated throughout the organization. The concept of reliability refers to the authority and trustworthiness of a record as a representation of the fact(s) it is about; that is, to its ability to stand for what it speaks of. In other words, reliability is the trustworthiness of a record's content. It can be inferred from two things: the degree of completeness of a record's documentary form and the degree of control exercised over the procedure (or workflow) in the course of which the record is generated. Reliability is then exclusively linked to a record's authorship and is the sole responsibility of the individual or organization that makes the record. Because, by definition, the content of a reliable record is trustworthy, and trustworthy content is, in turn, predicated on accurate data, it follows that a reliable record is also an accurate record.

An accurate record is one that contains correct, precise and exact data. Accuracy of a record may also indicate the absoluteness of the data it reports or its perfect or exclusive pertinence to the matter in question. The accuracy of a record is assumed when the record is created and used in the course of business processes to carry out business functions, based on the assumption that inaccurate records harm business interests. However, when records are transmitted across systems, refreshed, converted or migrated for continuous use, or the technology in which the record resides is upgraded, the data contained in the record must be verified to ensure their accuracy was not harmed by technical or human errors occurring in the transmission or transformation processes. The accuracy of the data must also be verified when records are created by importing data from other records systems. This verification of accuracy is the responsibility of the physical or juridical person receiving the data; however, such person is not responsible for the correctness of the data value, for which the sending person is accountable. Thus, the receiving person should issue a disclaimer regarding accuracy of records using other persons' data.

The concept of authenticity refers to the fact that a record is what it purports to be and has not been tampered with or otherwise corrupted. In other words, authenticity is the trustworthiness of a record as a record. An authentic record is as reliable and accurate as it was when first generated. Authenticity depends upon the record's transmission and the manner of its maintenance and custody. Authenticity is maintained and verifiable by maintaining the identity and integrity of a record. The identity of a record is established and maintained by indicating at a minimum the names of the persons participating in the creation of the record (e.g., author, addressee); the action or matter to which the record pertains; the date(s) of compilation, filing or transmission; the record's documentary form; the record's digital presentation (or format); the relationship of the record to other records through a classification code or a naming convention; and the existence of attachments. The integrity of a record is established and maintained by identifying the responsibility for the record through time by naming the handling person or office(s)[13] and the trusted records officer[14] or the recordkeeping office,[15] identifying access privileges[16] and access restrictions[17] and indicating any annotations or any modifications (technical or otherwise) made to the record by the persons having access to it.

Thus, record reliability is a quality that is established when a record is created and implies accuracy of the data contained in the record, while record accuracy and authenticity are qualities that are connected with the transmission and maintenance of the record. The latter are therefore the responsibility of both the records creator and any legitimate successor. Authenticity is protected and guaranteed through the adoption of methods that ensure the record is not manipulated, altered, or otherwise falsified after its creation, either during its transmission or in the course of its handling and preservation, within the recordkeeping system.[18]

---

[13] Handling office (or person) is defined as "The office (or officer) formally competent for carrying out the action to which the record relates or for the matter to which the record pertains" (InterPARES 2 Terminology Database, op. cit.).

[14] A trusted records officer (also called records keeper or records manager) is defined as "an individual or a unit within the creating organization who is responsible for keeping and managing the creator's records, who has no reason to alter the kept records or allow others to alter them and who is capable of implementing all of the benchmark requirements for authentic records" (Ibid.).

[15] Recordkeeping office is defined as "The office given the formal competence for designing, implementing and maintaining the creator's trusted recordkeeping system" (Ibid.).

[16] Access privileges is defined as "The authority to access a system to compile, classify, register, retrieve, annotate, read, transfer or destroy records, granted to a person, position or office within an organization or agency" (Ibid.).

[17] Access restrictions is defined as "The authority to read a record, granted to a person, position or office within an organization or agency" (Ibid.).

[18] See MacNeil et al., "Authenticity Task Force Report," op. cit.

(C5) A trusted record-making system should be used to generate records that can be presumed reliable.[19]

A trusted record-making system consists of a set of rules governing the making of records and a set of tools and mechanisms used to implement these rules. To generate reliable records, every record-making system should include in its design integrated business and documentary procedures, record metadata schemes, records forms, record-making access privileges and record-making technological requirements.

Integrated business and documentary procedures are business procedures linked to documentation procedures and to the classification system (i.e., the file management plan or taxonomy) established in the organization. This integration reinforces the control over record-making procedures: it supports the reliability of records by explicitly connecting records to the activities in which they participate and to the records organization system, thereby standardizing the procedures for creating and managing those records. The integration of business and documentary procedures also establishes the basis and central means to demonstrate ownership of and responsibility for the records. A record-making metadata scheme is a list of all metadata elements that need to be documented in the course of record-making processes for the purposes of uniquely identifying each record and enabling the maintenance of its integrity and the presumption of its authenticity. Such a scheme can also be used later to verify authenticity when questioned. Records forms are specifications of the documentary forms for the various types of records generated in the record-making system. Access privileges refer to the authority to compile, edit, annotate, read, retrieve, transfer and/or destroy records in the record-making system, granted to officers and employees by the records creator on the basis of position duties and business needs. Access privileges control access to the record-making system and are established in the course of integrating business and documentary procedures through connecting specific classes of records to the office of primary responsibility for a business function or activity. The establishment and implementation of access privileges is the most important step towards ensuring that the reliability of records can be presumed. Record-making technological requirements include the hardware and software specifications for the record-making system that have a direct impact on the documentary form of records.

(C6) A trusted recordkeeping system should be used to maintain records that can be presumed accurate and authentic. (P11, P12)

A trusted recordkeeping system consists of a set of rules governing the keeping of records and a set of tools and mechanisms used to implement these rules. Every recordkeeping system should include in its design a recordkeeping metadata scheme, a classification scheme, a retention schedule, a registration system, a recordkeeping retrieval system, recordkeeping technological requirements, recordkeeping access privileges and procedures for maintaining accurate and authentic records.

A recordkeeping metadata scheme is the list of all necessary metadata to be attached to each record to ensure its continuing identity and integrity in the recordkeeping system. A classification scheme is a plan for the systematic identification and arrangement of business activities and related records into categories according to logically structured conventions, methods and procedural rules. A retention schedule is a document specifying and authorizing the

[19] There is no corresponding Preserver Principle.

disposition of aggregations of records as identified in the classification scheme. A registration system is a method for assigning a unique identifier to each created record, linked to its identity and integrity metadata. Recordkeeping access privileges refer to the authority to classify, annotate, read, retrieve, transfer and/or destroy records in the recordkeeping system, granted to officers and employees by the records creator based on position duties and business needs. Typically, access to records for purposes of classification, transfer and destruction is given only to the trusted records officer of the organization. A recordkeeping retrieval system is a set of rules governing the searching and finding of records and/or information about records in a recordkeeping system and the tools and mechanisms used to implement these rules. Recordkeeping technological requirements include the hardware and software specifications for the recordkeeping system. The procedures for maintaining accurate and authentic records are the procedures designed to ensure that the data in the records and the identity and integrity of the records in the recordkeeping system are protected from accidental or malicious corruption or loss.

To improve efficiency and reduce the potential for human-induced error, the record-making and recordkeeping systems should be designed to automate, as much as possible, the creation of the identity and integrity metadata both at the point of records creation or modification (e.g., when migrated to a new system or file format), and whenever the aggregations to which the records belong are created or modified—every record unit should automatically inherit the metadata of the higher level in the classification at the point of creation as well as whenever there are updates to the metadata of the higher level.

A records creator should indicate in its records management policy that it is the trusted records officer's responsibility to manage the recordkeeping system. The role of the trusted records officer is analogous to that of a trusted custodian; thus, the trusted records officer should have the qualifications for a trusted custodian as stated in principle C8.

A recordkeeping system that complies with the above requirements and procedures in its design and management is capable of ensuring the accuracy and authenticity of records after their creation, since these requirements and procedures establish the maximum degree of control with regard to the maintenance and use of the records.

(C7) Preservation considerations should be embedded in all activities involved in record creation and maintenance if a creator wishes to maintain and preserve accurate and authentic records beyond its operational business needs. (P7)

The concept of the records lifecycle in archival science refers to the theory that records go through distinct phases, including creation, use and maintenance and disposition (i.e., destruction or permanent preservation).

It is essential for records creators dealing with records in digital form to understand that, differently from what is the case with traditional records, preservation is a continuous process that begins with the creation of the records. Traditionally, records are appraised for preservation at the disposition stage, when they are no longer needed for business purposes. With digital records, decisions regarding preservation must be made as close as possible to the creation stage because of the ease with which they can be manipulated and deleted or lost to technological obsolescence.

The notion that records preservation starts at the creation stage requires that preservation considerations be incorporated and manifested in the design of record-making and recordkeeping systems. Each aggregation of records appraised for preservation should be identified in accordance with the classification scheme and records retention schedule established by the records creator, and this identification should be indicated among the records metadata. The aggregations of records so identified should be monitored throughout their lifecycle so that appraisal decisions and preservation considerations can be updated and/or modified to accommodate any possible change occurring after they are first made. To monitor and implement appraisal decisions and preservation considerations, the designated preserver should be given access to the organization's recordkeeping system. Policies and procedures should be established to facilitate constant interaction between the records creator and its designated preserver.

(C8) A trusted custodian should be designated as the preserver of the creator's records. (P1)

The designated records preserver is the entity responsible for taking physical and legal custody of and preserving[20] (i.e., protecting and ensuring continuous access to) a creator's inactive records.[21] Be it an outside organization or an in-house unit, the role of the designated preserver should be that of a *trusted custodian* for a creator's records. To be considered as a trusted custodian, the preserver must:

- act as a neutral third party; that is, demonstrate that it has no stake in the content of the records and no reason to alter records under its custody and that it will not allow anybody to alter the records either accidentally or on purpose;
- be equipped with the knowledge and skills necessary to fulfil its responsibilities, which should be acquired through formal education in records and archives administration; and
- establish a trusted preservation system that is capable of ensuring that accurate and authentic copies of the creator's records are acquired and preserved.

For as long as the records are maintained by the creator in its recordkeeping system, they are active or semi-active records,[22] although under the responsibility of a trusted records officer. A records custodian trusted by the records creator as its designated preserver should maintain records that have been removed from the recordkeeping system for long-term or indefinite preservation. This trusted custodian will establish and maintain a preservation system to receive and preserve the creator's digital records. This involves ensuring that the accuracy and authenticity of the records received from the creator are assessed and maintained. Within the context of the preservation system, the designated preserver identifies appropriate preservation strategies and procedures, drawing on expertise from various disciplines, including archival science, computer science and law. The preservation procedures are implemented within the preservation system.

Only preservers that satisfy the requirements for trusted custodian are capable of fulfilling their duties of preserving authentic records over time and enabling a presumption of authenticity of the authentic copies they make for preservation purposes.

[20] The term "preservation" is defined as "The whole of the principles, policies, rules and strategies aimed at prolonging the existence of an object by maintaining it in a condition suitable for use, either in its original format or in a more persistent format, while leaving intact the object's intellectual form" (InterPARES 2 Terminology Database, op. cit.).

[21] An inactive record is defined as "A record that is no longer used in the day-to-day course of business, but which may be kept and occasionally used for legal, historical or operational purposes" (Ibid.).

[22] An active record is defined as "A record needed by the creator for the purpose of carrying out the actions for which it was created or for frequent reference" (Ibid). A semiactive record is defined as "A record that is no longer needed for the purpose of carrying out the action for which it was created, but which is needed by the records creator for reference" (Ibid.).

(C9) All business processes that contribute to the creation and/or use of the same records should be explicitly documented. (P10)

Records created in the course of carrying out one business function or one business process are often also used in the course of conducting other business functions or processes. In cases like this, records used in separate activities may be associated only with one activity in the records creator's record-making or recordkeeping system, or with none in some central "information" system or application. This practice creates difficulties for the records creator in identifying aggregations of records for accountability purposes and for its designated preserver in conducting appraisal and preservation activities.

It is recommended that policies and procedures be established that require detailed documentation of all business functions and processes contributing to the creation and use of the same records in any records creator's application or system and an explicit linkage between each record and the related workflow. Procedural manuals with such descriptions are effective in increasing the awareness of the impact of record-making and recordkeeping on the management of an organization. A subsequent different use of records after their creation can be captured by metadata, which are also capable of tracing the contexts in which records are generated.

(C10) Third-party intellectual property rights attached to the creator's records should be explicitly identified and managed in the record-making and recordkeeping systems. (P8)

Every records creator is usually aware that the records that it creates, or which are under its control or custody, contain information covered by intellectual property legislation. However, creators should also be aware that in some cases the intellectual property rights linked to a record may belong to a party other than the author and addressee.

All intellectual property rights attached to a record need to be documented in the metadata accompanying such record at the time that it is made or received and set aside. Intellectual property issues can significantly influence the reproduction of records, which is central to the processes of refreshing, converting and migrating records for either continuous use or preservation purposes. Subject to variations among different legislative environments, reproductions of records with intellectual property rights held by third parties may violate legislation that protects such rights. These issues must be identified and addressed at the stage of designing the record-making and recordkeeping systems. In the case of records identified for long-term preservation, long-term clearance of such rights should be addressed explicitly in the creator's record policy.

(C11) Privacy rights and obligations attached to the creator's records should be explicitly identified and protected in the record-making and recordkeeping systems. (P9)

Privacy legislation protects the rights of individuals with reference to personal data that may be part of any record used and maintained by a records creator with whom they have interacted. The limits of privacy depend on the legislative framework in which the records creator operates. The framework may be in conflict with the access policy linked to the mandate of the records creator and even with the access to information legislation in the same jurisdiction.

The presence of personal information within the records should be identified and documented within the metadata schema linked to the records in the record-making and recordkeeping systems of the creator. Metadata schemas that note and administer the use of personal information contained within the records must be embedded in record-making and recordkeeping systems. This will enable the protection of personal information through the establishment of system-wide access privileges. In cases where records are to be preserved indefinitely, privacy issues relating to access to records must be expressly resolved (i.e., explicit permissions must be sought from the individuals concerned), ideally prior to record creation. This is the best way to ensure that the records are managed in accordance with privacy legislation and that the preserver will be able to effectively include the privacy issues relevant to the records in the preservation feasibility study during appraisal. The designated preserver for each records creator should, as a trusted custodian, be granted access to records containing personal information to perform preservation activities. Processing of personal information for maintenance or preservation purposes is different from the use of it for research or business purposes. Regardless of the legislative framework, the records creator should be able to demonstrate that processing of records containing personal information does not put such information at risk of unauthorized access.

Responsibility for processing records containing personal data for maintenance and preservation purposes must reside with the records creator and its legitimate successors. Although the practice of outsourcing these functions to specialized commercial operators is authorized and regulated under most existing privacy legislation, the practice should still be avoided whenever possible to minimize the number of individuals authorized to access and/or process the records, thus reducing the risk of unauthorized disclosure of personal information in the records and of jeopardizing the ability to obtain permission to process personal information for maintenance or preservation purposes.

In the case of records that are not yet designated for permanent preservation, appraisal decisions should be taken before the initial mandate for processing personal information has expired to ensure that the legal basis for retaining such records is still in force.

(C12) Procedures for sharing records across different jurisdictions should be established on the basis of the legal requirements under which the records are created. (P13)

Records creators with branches in geographically separate areas (i.e., areas that are covered by different legislation), must be aware that different access, privacy and intellectual property laws may have an impact on their records-sharing activities. Such sharing activities encompass records exchange within the records creator or with outside organizations, such as governments or business partners. This includes providing records to a trusted preserver, where the latter operates in a legal environment different from that of the records creator.

The fact that records are freely accessible in one jurisdiction does not imply that they can be accessed in the same way in other jurisdictions. Records creators must investigate such issues and address them in their policies.

(C13) Reproductions of a record made by the creator in its usual and ordinary course of business and for its purposes and use, as part of its recordkeeping activities, have the same effects as the first manifestation, and each is to be considered at any given time the record of the creator. (P3)

In the digital environment, the first manifestation of a record, be it a draft, an original or a copy, only exists when first composed in the creator's record-making system, if it is an internal record, or when first received in the creator's recordkeeping system, if it is transmitted from the outside. When the record is closed and saved into the record-making or recordkeeping system, its first manifestation technically disappears, as the saving action decomposes it into its digital components. Any later manifestation of the digital record is a reproduction resulting from an assembly of its digital components. Conceptually, however, records creators can use any reproduction of a record's first manifestation as if it were the record's first manifestation, as long as the reproduction is made in the usual and ordinary course of carrying out business activities and used for such activities. This means that each reproduction in sequence should have the same admissibility in court as the record's first manifestation and be given the same weight.

To establish that a record is reproduced in the usual and ordinary course of business, it is necessary to set out routine procedures in writing. In effect, if reliable records have been generated in a trusted record-making system and their accuracy and authenticity have been maintained together with that of the received records in the creator's recordkeeping system, then all records should have the same authority and effects as their first manifestation.

Although, according to the theory of the record (i.e., diplomatics), an "original" record in a digital system is the first manifestation of a received record and, if after closing such manifestation the original no longer exists, it might be useful to look at three examples of statutory laws pertaining to the meaning of "original." Common to all three variations is the principle that it is the relationship of a record to the business of the creator that determines whether the record in question has the authority and effects of an original.

> Example 1: The U.S. Federal *Rules of Evidence* distinguishes between originals and duplicates, with greater value as evidence given to originals. For digital records, it is noteworthy that if "data are stored in a computer or similar device, any printout or other output readable by sight, shown to reflect the data accurately, is an 'original.'"[23]

> Example 2: The quality of being original is acknowledged in Italian legislation in terms of adding weight or greater trustworthiness to records. Italian legislation emphasizes the difference between digital data (original) and any kind of output of those data (copy), by establishing that "any data or document electronically created by any public administration represents a primary and original source of information that may be used to make copies on any kind of medium for all legal purposes."[24]

> Example 3: The *Electronic Signatures Law of the People's Republic of China* regards a digital record as an original if it meets the two following qualifications: it must be 1) capable of presenting the content effectively and of being retrieved and consulted at any moment, and 2) capable of unfailingly showing the integrity of the content from the moment of its completion. However, annotations made to a data electronic document [digital record] and changes of presentation occurring in the process of data exchanging, storing and displaying are not considered to affect its integrity.[25]

---

[23] United States House of Representatives, *Federal Rules of Evidence*, Article X. Contents of Writings, Recordings, and Photographs: Rule 1001. Definitions, Committee on the Judiciary, Committee Print No. 8 (December 31, 2004). Available at http://judiciary.house.gov/media/pdfs/printers/108th/evid2004.pdf. The same rule generalizes that "any counterpart" to the writing or recording "intended to have the same effect by a person executing or issuing it" is an original.
[24] Italy, DPR 445/2000, art. 9, par. 1. Available at http://www.parlamento.it/parlam/leggi/deleghe/00443dla.htm.
[25] China, *Electronic Signatures Law of the People's Republic of China*, art. 5. Translated by Sherry Xie. See also Sherry Xie (2005). "InterPARES 2 Project - Policy Cross-domain: Supplements to the Study of Archival Legislation in China (Report I)," 3. Available at http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CHINA_SUPPLEMENT.pdf.

Principles for Records Preservers

(P1) A designated records preserver fulfils the role of trusted custodian. (C8)

The designated records preserver is the entity responsible for taking physical and legal custody of and preserving (i.e., protecting and ensuring continuous access to) a creator's inactive records. Be it an outside organization or an in-house unit, the role of the designated preserver should be that of a *trusted custodian* for a creator's records. To be considered as a trusted custodian, the preserver must:

- act as a neutral third party; that is, demonstrate that it has no stake in the content of the records and no reason to alter records under its custody and that it will not allow anybody to alter the records either accidentally or on purpose;
- be equipped with the knowledge and skills necessary to fulfil its responsibilities, which should be acquired through formal education in records and archives administration; and
- establish a trusted preservation system that is capable of ensuring that accurate and authentic copies of the creator's records are acquired and preserved.

The acquisition of a creator's records is undertaken by the preserver, who, after having assessed the accuracy and authenticity of the records, produces an authentic copy of them from the creator's recordkeeping system. Records that are acquired this way are authentic copies of the records of the creator identified for long-term preservation, because they are made by the designated preserver in its role of trusted custodian.

The authentic copies of the creator's records are then kept by the trusted custodian in a trusted preservation system, which should include in its design a description and a retrieval system. This trusted preservation system must also have in place rules and procedures for the ongoing production of authentic copies as the existing system becomes obsolete and the technology is upgraded. This requirement is consistent with the final recommendations of InterPARES 1, which developed the *Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records*,[26] a set of requirements to be implemented by the preserver. It should be noted that the simple fact of reproducing records in the preserver's preservation system does not make the results authentic copies; such designation must be provided by the preserver's authority.

A sustainable preservation strategy requires close collaboration between a records creator and its designated preserver as trusted custodian. It is the preserver's responsibility to take the initiative in collaborating with the creator to establish acquisition and preservation procedures and in advising the creator in any records management activities essential to the preserver's acquisition and preservation activities.

(P2) Records preservation policies, strategies and standards should address the issues of record accuracy and authenticity expressly and separately. (C4)

An accurate record is one that contains correct, precise and exact data. The accuracy of a record is assumed when the record is created and used in the course of business processes to carry out business functions, based on the assumption that inaccurate records harm business interests. However, when records are transmitted across systems, refreshed, converted or migrated for preservation purposes, or the technology in which the record resides is upgraded,

[26] See MacNeil et al., "Authenticity Task Force Report," op. cit., and, more specifically, Authenticity Task Force, "Appendix 2."

the data contained in the record must be verified to ensure their accuracy was not harmed by technical or human errors occurring in the transmission or transformation processes. This verification of accuracy is the responsibility of the preserver who carries out the transmission or transformation process; however, such person is not responsible for the correctness of the data value, for which the creator remains accountable, just as is the case for the reliability of the records containing the data.

The concept of authenticity refers to the fact that a record is what it purports to be and has not been tampered with or otherwise corrupted. In other words, authenticity is the trustworthiness of a record as a record. A record is authentic if it can be demonstrated that it is as it was when created. An authentic record is as reliable and accurate as it was when first generated. Authenticity depends upon the record transmission and the manner of its preservation and custody. Thus, it is a responsibility of both the records creator and its legitimate successor (i.e., either the person or organization acquiring the function(s) from which the records in question result and the records themselves, or a designated records preserver).

Authenticity is protected and is verifiable by ensuring that the identity and the integrity of a record are maintained. The identity of a record is what distinguishes it from all other records. It is declared at the moment of creation by indicating at a minimum the following attributes: the names of the persons participating in the creation of the record (e.g., author, addressee); the action or matter to which the record pertains; the date(s) of compilation, filing or transmission; the record's documentary form; the record's digital presentation (or format); the relationship of the record to other records through a classification code or a naming convention; and the existence of attachments. The record identity so declared must be maintained intact through time first by the creator and its trusted records officer while the record is in active or semi-active use, and subsequently by the designated records preserver when the record is designated as inactive. The integrity of a record is its wholeness and soundness and can only be inferred from circumstantial evidence related to the person who held responsibility for the record through time, from access privileges and access restrictions and from the indication of any annotation or modification (technical or otherwise) that such person(s) with access to record might have made to it. Thus, the establishment and maintenance of record integrity are supported by declaring the following record attributes: the names of the handling office(s), the office of primary responsibility[27] for the record over time and/or the recordkeeping office and the designated preserver; the access privileges code[28] and the access restriction code;[29] and the list of annotations[30] and of format changes.[31]

Authenticity is not a quality that can be bestowed on records after their creation and maintenance by any preservation process. A preserver can only protect and maintain what was transferred under its responsibility. Authenticity is protected and maintained through the adoption of methods that ensure that the record is not manipulated, altered, or otherwise falsified after its transfer. It is the preserver's responsibility to assess the authenticity of records considered for acquisition into a preservation system and to ensure that it remains intact after the

[27] Office of primary responsibility is defined as "The office given the formal competence for maintaining the authoritative version or copy of records belonging to a given class within a classification scheme" (InterPARES 2 Terminology Database, op. cit.).
[28] Access privileges code is defined as "The indication of the person, position or office authorized to annotate a record, delete it, or remove it from the system" (Ibid.).
[29] Access restriction code is defined as "The indication of the person, position or office authorized to read a record" (Ibid.).
[30] List of annotations is defined as "Recorded information about additions made to a record after it has been created" (Ibid.).
[31] List of format changes is defined as "Recorded Information about modifications to a record's documentary form or digital format after it has been created" (Ibid.).

transfer to such system by respecting within the preserving unit or organization the same *Benchmark Requirements* that bind the creator (e.g., access privileges, measure against corruption or loss) and the *Baseline Requirements* for preservers.

(P3) Reproductions of a creator's records made for purposes of preservation by their trusted custodian are to be considered authentic copies of the creator's records. (C13)

Reproductions of digital records in the creator's record-making and recordkeeping systems made in the usual and ordinary course of activity for either action or reference purposes can be considered to have the same authority and effects as the first manifestation of the same records. Reproductions of a creator's records for preservation purposes rather than in response to a creator's business need are considered authentic copies of the records of the creator, because they are never used in their present manifestation for action or reference by the creator itself. The creator's records and their authentic preservation copies are the same records but at different phases in their lifecycle and thus at a different status of transmission.[32] The former are used by their creator to achieve business goals, while the latter are made by the preservers for preservation purposes.

Copies of records in the preserver's preservation system may not be designated authentic if the preserver has made them for purposes other than preservation; for example, a copy from which personal identifiers are removed may be made for access purposes. Ultimately, only the preserver has the authority to designate a copy as authentic.

(P4) Records preservation procedures should ensure that the digital components of records can be separately preserved and reassembled over time. (C2)

Every digital record is composed of one or more digital components. A digital component is a digital object that is part of one or more digital records, including any metadata necessary to order, structure or manifest content and that requires a given preservation action. For example, an e-mail that includes a picture and a digital signature will have at least four digital components (the header, the text, the picture and the digital signature). Reports with attachments in different formats will consist of more than one digital component, whereas a report with its attachments saved in one PDF file will consist of only one digital component. Although digital components are each stored separately, each digital component exists in a specific relationship to the other digital components that make up the record.

Preservation of digital records requires that all the digital components of a record be consistently identified, linked and stored in a way that they can be retrieved and reconstituted into a record having the same presentation it manifested when last closed. Each digital component requires one or more specific methods for decoding the bitstream and for presenting it for use over time. The bitstream can be altered, as a result of conversion, for example, as long as it continues to be able to fulfil its original role in the reproduction of the record. All digital components must be able to work together after they are altered; therefore, all changes need to be assessed by the preserver for the effects they may have on the record.

[32] In diplomatics, the status of transmission is the degree of perfection of record. There are three possible statuses of transmission: draft, original and copy. Copies are than further categorized according to their authority, and the most authoritative among the copies is the authentic copy; that is, a reproduction that is declared conforming to the reproduced entity by an officer having the authority to do so. Professional archivists are among such officers.

The preserver must be prepared to advise the creator, directly or through development of recommended standards, on the types of digital components that the preserver's system is able to sustain. Where standards governing the types and formats of digital components are common to both the record-making and recordkeeping systems and the record preservation system, the preserver can directly influence the creator towards those standards that will facilitate meeting the preservation requirements. Where no common standards exist or can reasonably be adopted, the preserver must understand the degree of interoperability of certain types and formats of digital components. This understanding will provide a basis for the preserver to assess the capability of the preservation system to preserve the digital components and their relationships as they emerge from the creator's record-making and recordkeeping systems.

Highly interoperable formats—that is, formats that are not tied to specific applications or versions of applications—are generally seen to provide a better basis for preservation work. It is important, however, not to focus exclusively on the interoperability of formats at the expense of the relationships between them that also must be preserved. For example, an HTML-based Web page may be comprised of digital components that are highly interoperable, but the version of HTML coding used to structure the components may be an old version with many deprecated terms (i.e., terms that are not recognized by current software browsers that may be used to reproduce the Web page).

(P5) Authentic copies should be made for preservation purposes only from the creator's records; that is, from digital objects that have a stable content and a fixed documentary form. (C1)

A record is defined by InterPARES, following the traditional archival definition, as "a document made or received in the course of a practical activity as an instrument or a by-product of such activity and set aside for action or reference."[33] This definition implies that, to be considered as a record, a digital object generated by the creator must first be a document; that is, must have stable content and fixed documentary form. Only digital objects possessing both are capable of serving the record's memorial function.

The concept of *stable content* is self-explanatory, as it simply refers to the fact that the data and the information in the record (i.e., the message the record is intended to convey) are unchanged and unchangeable. This implies that data or information cannot be overwritten, altered, deleted or added to. Thus, if one has a system that contains fluid, ever-changing data or information, one has no records in such a system until one decides to make one and to save it with its unalterable content.

The concept of *fixed form* is more complex. A digital object has a fixed form when its binary content is stored so that the message it conveys can be rendered with the same documentary presentation it had on the screen when first saved. Because the same documentary presentation of a record can be produced by a variety of digital presentations, fixed form does not imply that the bitstreams must remain intact over time. It is possible to change the way a record is contained in a computer file without changing the record; for example, if a digital object generated in '.doc' format is later saved in '.pdf' format, the way it manifests itself on the screen—its documentary presentation, or "documentary form"—has not changed, so one can say that the object has a fixed form.

[33] See the InterPARES 2 Terminology Database, op. cit.

One can also produce digital information that can take several different documentary forms. This means that the same content can be presented on the screen in several different ways, the various types of graphs available in spreadsheet software being one example. In this case, each presentation of such a digital object in the limited series of possibilities allowed by the system is to be considered as a different view of the same record having stable content and fixed form.

In addition, one has to consider the concept of "bounded variability,"[34] which refers to changes to the form and/or content of a digital record that are limited and controlled by fixed rules, so that the same query, request or interaction always generates the same result. In such cases, variations in the record's form and content are either caused by technology, such as different operating systems or applications used to access the document, or by the intention of the author or writer of the document. Where content is concerned, while, as mentioned, the same query will always return the same subset, its presentation might vary within an allowed range, such as image magnification. In consideration of the fact that what causes these variations also limits them, they are not considered to be violations of the requirements of stable content and fixed form.

Based on this understanding, any preservation policy should clearly state that reproductions of authentic copies for preservation purposes can only be made from the creator's records, as identified by the creator.[35]

The preserver should know (or help establish) the creator's criteria for identifying the digital objects that are maintained as records and the methods employed to stabilize their content and fix their form. This is consistent with the preserver's responsibility to advise the creator on its record creation processes and technologies. This advising activity will also provide the preserver with the critical information needed to understand the business activities and processes that caused the records to come into being and with the ability to assess their continuing identity and integrity.

(P6) Preservation requirements should be articulated in terms of the purpose or desired outcome of preservation, rather than in terms of the specific technologies available. (C3)

Digital records rely, by definition, on computer technology, and any instance of a record exists within a specific technological environment. For this reason, it may seem useful to establish record preservation requirements in terms of the technological characteristics of the records or the technological applications in which the records may reside. However, not only do technologies change, sometimes very frequently, but they also are governed by proprietary considerations established and modified at will by their developers. Both these factors can significantly affect the continued accessibility of digital records over time. For these reasons, references to specific technologies should not be included in preservation policies and standards. Only the requirements and obligations that the records are designed to support should be explicit within record preservation policies and standards. It is only at the level of implementation that specific technologies should, indeed must, be named.

Technological solutions to record preservation issues are dynamic, meaning that they will evolve as the technology evolves. This affects record preservation in two ways. First, it makes it possible to adopt new strategies to meet preservation needs, as happened with the use of XML to support the long-term preservation of structured records. Second, it creates opportunities for drawing on expertise from a number of disciplines. These two issues are interconnected. Thus,

[34] See Duranti and Thibodeau, "The Concept of Record," op. cit.
[35] See principle C1 in the Principles for Creators regarding the identification of records.

for example, while utilization of XML is, by itself, only one activity for preservation, it might be matched with using data grid technology as a stable and enduring platform to support XML-based records. By experimenting with these combinations, new archival knowledge will continue to be both acquired and required.

Technological solutions also need to be specific to be effective. Although the general theory and methodology of digital preservation applies to all digital records, the preservation solutions for different types of records require different methods. These should be based on the specific context in which the records are created and maintained, the functions and activities to which the records are linked and the technologies employed for record-making and recordkeeping to ensure the best solutions are designed for preserving each type of record.

Preservation policies that are expressed in terms of record requirements rather than technologies will be more stable, needing updates only if the record requirements change, rather than as the technology changes. Preservation action plans will likely need to be updated more frequently to identify appropriate technological solutions for the digital preservation of specific aggregations of records. The identified solutions must be monitored with regard to the possible need for modifying and updating.

(P7) Preservation considerations should be embedded in all activities involved in each phase of the records lifecycle if their continuing authentic existence over the long term is to be ensured. (C7)

The concept of the records lifecycle in archival science refers to the theory that records go through distinct phases, including creation, use and maintenance and disposition (destruction or permanent preservation).

It is essential for preservers who acquire digital records to understand that, differently from what is the case with traditional records, preservation is a continuous process that begins with the creation of the records. Analogue records are appraised for preservation at the disposition stage, when they are no longer needed by the creator for business purposes. With digital records, decisions relevant to preservation must be made as close as possible to the creation stage because of the ease and the speed with which digital objects can be manipulated, deleted by accident or on purpose, or lost to technological obsolescence.

The notion that records preservation starts at the creation stage requires that preservation considerations be incorporated and manifested in the design of record-making and recordkeeping systems. Each aggregation of records appraised for preservation should be identified in accordance with the classification scheme and the records retention schedule established by the records creator in collaboration with the preserver, and this identification should be indicated in the records metadata. The records so identified should be monitored throughout their lifecycle by the preserver, so that appraisal decisions and preservation considerations can be updated to accommodate any possible changes occurring after they are first made. Appraisal decisions need to be reviewed to ensure that the information about the appraised records is still valid, that changes to the records and their context have not adversely affected their identity or integrity and that the details of the process of carrying out disposition are still workable and applicable to the records. To monitor and implement appraisal decisions and preservation considerations, the designated preserver should obtain continuing access to the records creator's recordkeeping system within limits agreed upon with the creator and reflected in the preserver's access privileges. The preserver should establish procedures to facilitate constant interaction with the records creator.

(P8) Third-party intellectual property rights attached to the creator's records should be explicitly identified and managed in the preservation system. (C10)

Preservers know that records under records creators' control usually contain information covered by intellectual property legislation. They should also be aware that, in some cases, the intellectual property rights attached to records belong to a party other than the author; that is, the intellectual property rights reside with a third party. Third-party intellectual property rights should be documented in the metadata accompanying such records because they influence the processes of refreshing, converting and migrating them for either continuous use or preservation purposes. Subject to variations in different legislative environments, reproductions of records with third-party intellectual property rights attached to them may violate legislation that protects such rights. In the case of records identified for long-term preservation, long-term clearance of such rights should be addressed explicitly with the records creator.

Because preservation in a digital environment involves making copies, intellectual property rights have become an issue, not just for access as in the past, but for preservation. It is the preserver's responsibility; first, to advise the creator on how to address intellectual property issues in its record-making and recordkeeping systems, and, second, to ensure that intellectual property issues are addressed in the design of the preservation system. In particular, any issues relevant to third-party intellectual property rights should be cleared before the transfer of records to be preserved from the creator to the preserver. The latter must consider these issues as a part of the assessment of feasibility of preservation.

(P9) Privacy rights and obligations attached to the creator's records should be explicitly identified and protected in the preservation system. (C11)

Privacy legislation protects the rights of individuals with reference to personal data that may be part of any record used and maintained by a records creator with whom they have interacted. The limits of privacy depend on the legislative framework in which the records creator operates. It may be in conflict with the access policy linked to the mandate of the records creator and even with the access to information legislation in the same jurisdiction. Besides lobbying for exceptions, the designated preserver should ensure that the consequences of the existing situation for preservation and access are clearly understood.

The presence of personal information within the records should be identified and documented among the metadata linked to the records in the record-making and recordkeeping systems of the creators. This is the best way to ensure that the records are managed in accordance with privacy legislation and that the preserver will be able to effectively include the privacy issues relevant to the records in the preservation feasibility study during appraisal. The designated preserver for each creator should, as a trusted custodian, obtain access to records containing personal information to perform preservation activities. Archival processing of personal information for preservation purposes is different from the use of it for research or business purposes. Regardless of the legislative framework, the creator and the preserver should be able to demonstrate that archival processing of records containing personal information does not put such information at risk of unauthorized access.

Preservers should also insist that responsibility for processing records containing personal data for preservation purposes must reside with the records creator and its legitimate successors. Although the practice of outsourcing these preservation functions to specialized commercial operators may be authorized and regulated under most existing privacy legislation, the practice should still be avoided whenever possible to minimize the number of individuals authorized to access and/or process the records, thus reducing the risk of unauthorized disclosure of personal information in the records and of jeopardizing the ability to obtain permission to process personal information for preservation purposes.

In the case of records that are not yet designated for permanent preservation, appraisal decisions should be taken before the initial mandate for processing personal information has expired to ensure that the legal basis for retaining such records is still in force.

(P10) Archival appraisal should identify and analyze all the business processes that contribute to the creation and/or use of the same records. (C9)

A record may be created for one purpose and then subsequently used for different purposes by different persons. Any appraisal decision should consider all uses of the record and be aware of the business processes behind them. This is necessary to make an informed decision about what to preserve as well as to be able to dispose effectively of all possible copies of the records that have not been selected for preservation.

The use of records or information within records by different business processes may be desirable from the creator's standpoint in terms of providing a degree of interoperability among the creator's information and record systems. In such situations, the preserver should advise the creator that metadata attached to records used by many business processes must identify each relevant business process. This is critical for the creator because it ensures the authenticity of the records by establishing their identity and integrity in each context. It is also critical for the preserver who must understand all contexts in which the records were used to effectively undertake appraisal and also to meet the baseline requirements for maintaining authenticity for any records acquired into the preservation system.

(P11) Archival appraisal should assess the authenticity of the records. (C6)

Appraisal decisions should be made by compiling information about kept records and their context(s), assessing their value and determining the feasibility of their preservation.[36]

As part of the assessment of value, preservers must establish the grounds for presuming that the records being appraised are authentic. This means that preservers must ensure that each record identity has been documented and maintained as documented and must ascertain the degree to which the records' creator has guaranteed their integrity by making sure that its records are intact and uncorrupted. The evidence supporting the presumption of authenticity must be measured against the InterPARES *Benchmark Requirements*.[37]

(P12) Archival description should be used as a collective authentication of the records in an archival fonds. (C6)

Archival description of a fonds emerges from the comprehensive analysis of the various relationships interwoven in the course of the formation and accumulation of records and therefore is the most reliable means of establishing the continued authenticity of a body of

[36] See Terry Eastwood et al., "Part Two – Choosing to Preserve: The Selection of Electronic Records: Appraisal Task Force Report," in Duranti, *Long-term Preservation*, op. cit., 67–98. Online reprint available at http://www.interpares.org/book/interpares_book_e_part2.pdf.

[37] See the already cited benchmark requirements in MacNeil et al., "Appraisal Task Force Report," op. cit.

interrelated records. While the authenticity of individual records can be in part established through their metadata, the authenticity of aggregations of records (i.e., file, series or fonds), can only be proved through archival description.

It has always been the function, either explicit or implicit, of archival description to authenticate the records by perpetuating their administrative and documentary relationships; but, with digital records, this function has moved to the forefront. In fact, as original digital records disappear and an interminable chain of non-identical reproductions follows them, the researchers looking at the last of those reproductions will not find in it any information regarding provenance, authority, context or authenticity.

The authentication function of archival description is different from that of a certificate of authenticity, because it is not simply an attestation of the authenticity of individual records, but a collective attestation of the authenticity of the records of a fonds and of all their interrelationships as made explicit by their administrative, custodial and technological history (including a description of the recordkeeping system(s) within which they have been maintained and used), the scope and content and the hierarchical representation of the records aggregates. It is also different both from the identity and integrity metadata attached to individual records, which are part of the record itself and are reproduced time after time with it and from the additional metadata attached to records aggregations (e.g., file, series) within the recordkeeping system to identify them and document their technological transformations.

The unique function of archival description is to provide an historical view of the records and of their becoming, while presenting them as a universality in which each member's individuality is subject to the bond of a common provenance and destination.

(P13) Procedures for providing access to records created in one jurisdiction to users in other jurisdictions should be established on the basis of the legal environment in which the records were created. (C13)

Different jurisdictions may have different laws and regulations with regard to access rights in relation to the protection of privacy, intellectual property and any other kind of public or private interests (e.g., market sensitive records). Preservers who are a unit of a records creator (e.g., in-house archival programs or archives) that has geographically separated branches falling under different legislation must be aware of the impact of such diverse legal contexts on their records-sharing activities. This will affect access policies relevant to both internal and external sharing activities.

APPENDIX 20

CREATOR GUIDELINES
Making and Maintaining Digital Materials: Guidelines for Individuals

by

Philip Eppard
University of Albany, State University of New York

CREATOR GUIDELINES
Making and Maintaining Digital Materials: Guidelines for Individuals[1]

Introduction

Most information today is created and stored in digital form. The advantages of the digital medium are by now familiar to everyone. Documents can be created quickly and edited and revised with ease. Thanks to the Internet, they can be distributed globally with lightning-like speed. They can be manipulated in ways that allow them to be used for multiple purposes. The digital medium also solves the longstanding storage problems associated with large files of paper records.

The blessings of the digital era, however, are not without their costs. Only in recent years have people begun to fully grasp the many problems inherent in the digital medium. For example, there is the fact that digital information can only be accessed using a computer. Furthermore, the computer must be equipped with the necessary software to be able to read the bit strings contained on the disc or tape. Ease of reproduction and the proliferation of copies make it more difficult to identify a complete or final version of a digital document. Easy distribution of information on the Internet makes the preservation of intellectual property rights difficult. Finally, all digital materials are vulnerable to viruses and simple technology failure, as well as to the rapid developments in software and hardware that risk making them inaccessible very quickly.

With all of these problems, it is little wonder that some people yearn for the comforting tangibility of paper. Yet although our systems for creating and maintaining information will likely continue for some time to be hybrid systems—that is, containing both paper and digital materials—there is clearly no turning back from the digital revolution. Consequently, everyone should be aware of the risks faced by digital materials and know how best to minimize these risks.

These guidelines have been developed for individuals who create digital materials in the course of their professional and personal activities to help them make informed decisions about making and maintaining these materials in ways that will help ensure their preservation for as long as they are needed. They may also be useful for small organizations or groups of individuals, such as medical offices, consulting groups or teams of research scientists.

Although these guidelines can be applied to various kinds of digital publications, documents and data, they are especially important for digital records. Records are the documents that you make, receive and use in your activities, and that you keep because you may need them later or because you want to have reliable evidence of what you have done. Therefore, you need to be especially careful in maintaining and preserving them. These guidelines are applicable to records that need to be maintained for only a short period of time as well as to those that require long-term maintenance. Adherence to these guidelines will help ensure that records that merit long-term preservation in an archival repository will be accessible when they are turned over to the care of a trusted custodian.

[1] These Guidelines have also been issued in an illustrated booklet form that is freely available at http://www.interpares.org/display_file.cfm?doc=ip2(pub)creator_guidelines_booklet.pdf.

Definitions

Before presenting recommendations to guide you in making and maintaining digital materials, it will be both necessary and helpful to clarify the meaning of some of the terms used in this document.

For the purposes of these guidelines, a *record* is defined as any document created (i.e., made or received and saved for further action or reference) by a physical or corporate person in the course of a practical activity as an instrument and by-product of that activity. A *publication* is defined as a document intended for dissemination or distribution to the public at large. All records and publications are documents and contain data. A *document* is information affixed to a medium in a fixed form; *information* is an assemblage of data intended for communication over time or space; and *data* are the smallest meaningful and indivisible pieces of information.

These guidelines aim at providing recommendations for the creation and maintenance of reliable digital materials in general, and records in particular, that can be accurately and authentically maintained and preserved over time. To facilitate their application, however, the terms “reliability,” “accuracy,” “authenticity” and “authentication” need to be defined.

For the purposes of these guidelines, *reliability* is the trustworthiness of digital materials as statements of fact or as content. It is the responsibility of the author of the materials, be that author an individual or the corporate person in whose name an individual is writing, and is assessed on the basis of the material’s completeness and accuracy and of the degree of control exercised on the process of its creation.

Accuracy is the degree to which the data in the materials are precise, correct, truthful and free of error or distortion. To ensure accuracy, one must exercise control on the processes of creation, transmission, maintenance and preservation of the materials. Over time, the responsibility for accuracy shifts from the author to the keeper of the materials and later to the long-term preserver of the materials (if applicable).

Authenticity refers to the fact that the materials are what they purport to be and have not been tampered with or otherwise corrupted. Thus, with respect to records in particular, authenticity refers to the trustworthiness of records as records. To ensure that authenticity can be presumed and maintained over time, one must define and maintain the identity of the materials and protect their integrity. Authenticity is at risk whenever materials are transmitted across space and time. Over time, the responsibility for authenticity moves from the keeper to the long-term preserver of the materials.

Authentication is a declaration of authenticity, resulting either from the insertion or the addition of elements or statements to the materials in question, and the rules governing it are established by legislation. Thus, it is a means of proving that materials are what they purport to be at a given moment in time. Digital authentication measures, like the use of digital signatures, only ensure that the materials are authentic when received and cannot be repudiated, but not that they will stay authentic afterwards.

Recommendations

1. Select hardware, software and file formats that offer the best hope for ensuring that digital materials will remain easily accessible over time.

Accessing digital materials depends on having the appropriate software. Software that is not compatible with previous versions (backward compatibility) or with future versions (forward compatibility) makes it difficult to access records over time. Software for one application also needs to work well with that of other applications and systems (interoperability). Paying attention to the following factors can help ensure that your software and hardware maintain accessibility.

A. *Choose software that presents materials as they originally appeared*. Ideally, materials should keep the same look over time to be fully intelligible and accessible. Be sure that new software will be able to read your older materials in the software format in which you kept it and display it on the screen in the same documentary form in which it was originally displayed. In other words, new software should be backward compatible with older software.

B. *Choose software and hardware that allow you to share digital materials easily*. Software should be able to accept and output files in a number of different formats. The ability to interact easily with other technology is called *interoperability*. It will make it easier to access your materials and also to move them to other systems.

C. *Use software that adheres to standards*. This is one of the best things you can do to ensure your material will last. Standards endorsed by national and international organizations are best. These are called *de jure* standards.[2] If these do not exist for your material, you can help ensure longevity by adopting software that is very widely used. In the absence of an official standard, such software is often referred to as a *de facto* standard.[3] Open source software; that is, freely available non-proprietary software, is preferable (see subsection G on the next page).

D. *Keep the specifications of software*. This kind of documentation (e.g. the owner's manuals or any other more detailed description of the software you might have) will be essential in the future to access the materials or to migrate them to a new computer environment as technology advances. It is particularly important to fully document any software that you build yourself.

E. *If you customize software, make sure you document the changes you make*. Give detailed information about the changes and describe clearly the characteristics and features of the material these changes produce, as well as the outcomes you are trying to achieve by

[2] Defined as: A standard adopted by an official standards-setting body, whether national (e.g., ANSI), multi-national (e.g., CEN) or international (e.g., ISO). For computer file formats, two recent de jure standards are PDF/A (PDF standard for archiving) and ODF (OASIS OpenDocument Format).
[3] Defined as: A standard not adopted by any official standards-setting body, but nevertheless widely used and recognized by its users as a standard. Well known and widely used computer file formats that are considered de facto standards include PDF, TIFF, DOC and ZIP.

customizing the software. A good way to do this is to include the information as comments in the software code. The information will not get lost, as it is part of the file, and it will be very helpful to those who need to make adjustments later, as technology advances.

F. *Document the construction of your system as a whole to help ensure its accessibility*. You should document your system's structure and functions. This means identifying its hardware and software components, including peripherals, its operating system and software packages. Such documentation will identify how the software packages represent information, and how they process it and communicate it to each other and to users. These basic specifications will ensure that those who come after you understand the context in which you are working now. They will provide the information necessary to update the system as hardware and software evolve.

G. *Choose widely-used, non-proprietary, platform-independent, uncompressed formats with freely available specifications where possible*. These are often called "open formats," which means that their specification is published and freely available. However, it may also mean that the format is free of patent or royalty fees or the possibility of such fees being applied in the future, and/or that it is widely adopted. It should be noted that "open" formats are not necessarily the same as formats produced by *open source software*, as the latter term describes software for which the code is made freely available and can be modified. Open source software does not always produce non-proprietary formats. Distinguish between file formats, wrapper (or container) formats and tagged formats such as XML-tagged files, and ensure that version, encoding and other characteristics are clear and fully specified. For XML files, make sure that the files are well-formed and valid and accompanied by the relevant DTDs or schemas. If it is not convenient for you to follow this recommendation, consult with an archives that accepts digital materials and choose among the formats that it recommends for long-term preservation. You should not compress your digital materials, if at all possible, since this can lead to problems for their long-term preservation. If you need to compress them, choose lossless compression techniques that conform to accepted international standards.

2. Ensure that digital materials maintained as records are stable and fixed both in their content and in their form.

One of the great advantages of digital materials is the ease with which information can be edited, revised or updated. But this also means that important information can be changed or even lost, accidentally or on purpose. This is a particularly important problem for records, because one of the characteristics of a record is that its content is unchanged and unchangeable. This implies that the information and the data in the record cannot be overwritten, altered, deleted or expanded. A system that contains fluid, ever-changing information or data does not really contain records until someone decides to make them and save them with *fixed form*[4] and *stable content*.[5]

[4] Defined as: The quality of a record that ensures the documentary appearance or presentation is the same each time the record is retrieved.
[5] Defined as: The quality of a record that makes the information and data contained in it immutable, and requires changes to be made by appending an update or creating a new version.

Although the idea of stable content is fairly simple, the concept of fixed form is more complex. Essentially, it means that the message conveyed by a digital record (or other digital object) can be rendered with the same documentary presentation it had on the screen when it was made or received and first saved. The bit streams that compose the digital record and determine its digital presentation (i.e., its file format) may change, but its documentary presentation must not change. A simple example is when a document created in Microsoft Word is later saved as an Adobe PDF file. Although the document's digital presentation has changed—from a Microsoft Word .doc file format to an Adobe .pdf file format—the documentary presentation of the document—also called its *documentary form*[6]—has not changed, and therefore we can say that the document has a fixed form.

In some cases, digital materials can be presented in several different ways—in other words, the information they convey can take different documentary forms. For example, statistical data can be presented as a pie chart, a bar chart or a table. However, the possible variations of these displays are usually limited by the system. In such cases, we can regard each documentary presentation as having stable content and fixed form, since the information is selected from a fixed store of data within the system and the system's rules govern the form of its documentary presentation(s).

A similar situation occurs when the selection of both content and form is from a large store of fixed information that is only partially accessed every time a user queries the system. If the same query always produces the same output as to content and documentary form, the output can be regarded as having stable content and fixed form. Thus, if you, as the author of the record, establish fixed rules for the selection of its content and of its documentary form that only allow for a known and stable range of variability— that is, endow it with *bounded variability*[7]—then you can claim that your material has stable content and fixed form.

The concern for the documentary presentation of digital materials is particularly important for maintaining and assessing the reliability and accuracy of records. Future upgrades, conversions or migrations of data may result in changes to the documentary form. Therefore, you would be wise to first establish the documentary form of records associated with each activity or procedure and then identify the essential characteristics (i.e., the essential *intrinsic* and *extrinsic* elements[8]) of each documentary presentation or form. This will help alert you to any changes in the future that would imply a loss of identity and integrity of the record, especially if you are active in the sphere of digital art, where a certified description of those essential characteristics by the artist would help support the recognition of the intellectual property rights linked to work so described.

[6] Defined as: The rules of representation according to which the content of a record, its administrative and documentary context and its authority are communicated. Documentary form possesses both extrinsic and intrinsic elements.

[7] Defined as: The quality of a record that ensures that its documentary presentations are limited and controlled by fixed rules and a stable store of content data, form data and composition data, so that the same user activity, query, request or interaction always generates the same result.

[8] *Intrinsic Elements* are defined as: The elements of a record that convey the action in which the record participates and its immediate context, including the names of the persons involved in its creation, the name and description of the action or matter to which it pertains, the date(s) of creation and transmission, etc. *Extrinsic Elements* are defined as: The elements of a record that constitute its external appearance, including presentation features such as font, graphics, images, sounds, layouts, hyperlinks, image resolutions, etc., as well as digital signatures, seals, and time stamps and special signs (digital watermarks, logos, crests, etc.).

3. Ensure that digital materials are properly identified.

Giving a meaningful name to a computer file helps identify its content and makes it easier to find. The full identification of records is more complex than just naming files, however. Full identification is essential in distinguishing records from each other, in distinguishing different versions of a single record and in providing evidence of the identity of a record from the moment of its creation through its long-term preservation.[9]

The information about digital materials that supports their identification and retrieval is commonly referred to as *metadata*.[10] Most software applications automatically tag all digital materials with some data about their identity because this information is necessary to locate documents effectively. Without metadata, it would be nearly impossible to find a document without opening and reading through a folder or several directories. Metadata describe the properties or attributes of digital materials. In the case of records, however, these properties or attributes are also necessary to maintain and assess their authenticity, and that is why it is important to ensure that all the essential ones are recorded and that they are correct.

The properties or attributes conveying the identity of digital materials are referred to as identity *metadata*.[11] These include:

a. *Names of the persons involved in the creation of the digital materials*. These include:
    - the *author*—the physical or corporate person(s) responsible for issuing the materials;
    - the *writer*—the physical person(s) or position(s) responsible for articulating the content of the materials;
    - the *originator*—the physical person, position or office responsible for the electronic account or technical environment where the materials are generated and/or from which it is transmitted;[12]
    - the *addressee*—the physical or corporate person(s) for whom the materials are intended; and
    - the *recipient*—the physical or corporate person(s) to whom the materials may be copied or blind copied.

b. *Name of the action or matter*—in other words, the title or subject.

c. *Documentary form*—in other words, whether it is a report, a letter, a contract, a table, a list, etc.

d. *Digital presentation*—in other words, format, wrapper, encoding, etc.

e. *Date(s) of creation and transmission*. These include:
    - the *chronological date* written on the materials or on which the materials were compiled;

[9] In this context, *identity* is defined as: The whole of the characteristics of a document or a record that uniquely identify it and distinguish it from any other document or record. With integrity, a component of authenticity. (See also Recommendation 4.)

[10] Defined as: Information that characterizes another information resource, especially for purposes of documenting, describing, preserving or managing that resource.

[11] Defined as: The properties or attributes conveying the identity of a digital object that is to be kept as a record. (See also Recommendation 5.)

[12] Identification of the originator is only important in cases where the person, position or office responsible for physically creating and/or transmitting the materials is neither the author nor the writer, and when the presence of the originator's name appearing on, or in association with, the materials calls into question the actual author and/or writer of the materials. This is most commonly associated with e-mails in instances where the name of the originator appears in the header of an e-mail and/or its attachments that were in fact authored and/or written by another person, but physically manifested and/or transmitted on behalf of that person by the originator.

- the *dates of transmission and/or receipt*; and
- the *archival or filing date*—in other words, the date when the materials were associated with a computer folder or directory, or other classification scheme or filing plan (see Recommendation 5).

f. *Expression of documentary context*—for example, a classification code, or the name of the computer folder or directory, or comparable filing unit within the classification scheme or filing plan to which the materials are associated, and the name of the broader group of records in which the materials belong (see also Recommendation 5).
g. *Indication of attachments*, if applicable.
h. *Indication of copyright or other intellectual rights*, if applicable.
i. *Indication of the presence or removal of a digital signature*, if applicable (see Recommendation 6, Technology-dependent Authentication section).
j. *Indication of other forms of authentication*, if applicable. This could include, for example, the presence of a corroboration (i.e., an explicit mention of the means used to validate the record); an attestation (i.e., the validation of a record by those who took part in the issuing of it, and by witnesses to the action or to the 'signing' of the record); a subscription (i.e., the name of the author or writer appearing at the bottom of the document), or a qualification of signature (i.e., the mention of the title, capacity and/or address of the person or persons signing the record).
k. *Indication of the draft or version number*, if applicable.
l. *Existence and location of duplicate materials outside the digital system*, if applicable. If multiple copies of a document exist, you should indicate which one is the official or *authoritative copy*.[13] If the document is certified by the author as an "approved reproduction" of a work (for example, a digital work of art), indication of the existence of such certification is required. If the document comprises material copyrighted by different author(s), indication of copyright clearance (or lack thereof) with related dates is necessary.

4. Ensure that digital materials carry information that will help verify their integrity.

Although the identity metadata help distinguish digital materials from one another, another set of metadata allows users to infer that the materials are the same as when they were created (although not to verify or demonstrate it, because this would require comparison with a copy of the materials kept elsewhere). These metadata can be referred to as *integrity metadata* (see below). Digital materials have *integrity*[14] if they are intact and uncorrupted, that is, if the messages that they are meant to communicate to achieve their purposes are unaltered. This means that the physical integrity of digital materials, such as the proper number of bit strings, may be compromised, provided that the articulation of the content and its required elements of *documentary form* (see Recommendation 2) remain the same. The content and the data in it are considered to be unaltered if they are identical as to the value and presentation (i.e., position on the screen) of the content and data in the first saved manifestation of the material. The attributes that relate to the integrity of digital materials have to do with the maintenance of the materials, including the responsibility for their proper handling, such as overseeing and documenting any

[13] Defined as: The instance of a record that is considered by the creator to be its official record and is usually subject to procedural controls that are not required for other instances.
[14] Defined as: The quality of being complete and unaltered in all essential respects. With identity, a component of authenticity.

technological transformations or transfers of the materials to other systems. The integrity metadata include:

a. *Names of handling person/office*—the person or office using the materials to carry out business.
b. *Name of person or office with primary responsibility for keeping the materials*—this may be the same as the handling person/office.
c. *Indication of annotations added to the materials*, if applicable.
d. *Indication of any technical changes to the materials or to the application(s) responsible for managing and providing access to the materials*—for example, change of encoding, wrapper or format, upgrading from one version to another of an application, conversion from several linked digital components to one component only (e.g., by embedding directly in the materials digital components that were previously only linked to the materials, such as audio, video, graphic or text elements like fonts).
e. *Access restriction code*—indication of the person, position or office authorized to read the materials, if applicable.
f. *Access privileges code*—indication of the person, position or office authorized to annotate the materials, delete them, or remove them from the system, if applicable.
g. *Vital record code*—indication of the degree of importance of the record to continue the activity for which it was created or the business of the person/office that created it, if applicable.[15]
h. *Planned disposition*—for example, removal from the live system to storage outside the system; transfer to the care of a *trusted custodian* (see Recommendation 10); scheduled deletion.

5. Organize digital materials into logical groupings.

The management and retrieval of your digital materials can be enhanced if you can handle them in large sets, rather than one by one. Therefore, it is important that you group your digital materials in some logical manner. The categories chosen may reflect the way you work, your activities, procedures, thematic areas, or some sort of structural organization. Separating your records from other digital materials is an important first step. The organization of your records may be based on the different types of records or the length of time for which certain kinds of records need to be kept. These groupings can be related to each other in a hierarchical or flat way, as best suits your needs. Generally, this structure should be consistent with the organization of any paper records you have (or records in other media), so that all records related to the same activity or subject, or of the same type, can be easily identified and retrieved as part of one conceptual grouping, as needed. Your organization scheme should be recorded in a document that shows all the groupings of materials, describes them in a brief sentence and indicates how they are related. In this document, which is called a *classification scheme*[16] or filing plan, each grouping of records can be assigned a code or a name that should be linked to each record belonging in the same grouping no matter what the medium or location: thus, the records assigned to each grouping will share such code or name, followed by a number that indicates

[15] The *vital record code* only pertains to specific communities of practices, such as legal and medical offices, who must identify the records that are vital to the continuance of their business in case of disaster and who would therefore exercise special protection measures on those records.

[16] Defined as: A plan for the systematic identification and arrangement of business activities and records into categories according to logically structured conventions, methods and procedural rules. (See also Recommendation 3.)

their sequence. This identifier should be recorded among the *identity metadata*[17] of your digital records and on the face of your paper records belonging to the same grouping and should be unique for each record.

Identifying how long groupings of records need to be retained will facilitate their management while they are regularly needed and help ensure that records that need or merit long-term preservation are tagged early and given proper protection to ensure their survival.

You will find it easier and more efficient to assign a retention period—the length of time you want or need to keep materials—to a grouping of materials, rather than to individual items. Trying to ensure that some things are kept as long as needed while weeding out things that are no longer needed is simply too cumbersome at the individual item level. Although you may think that within a grouping some records should be kept longer than others, not only will you save time if you keep the whole grouping, but you will also have more complete information when you need to refer to the records. However, for some types of records, you can create subgroups within each given grouping on the basis of the retention period.

6. Use authentication techniques that foster the maintenance and preservation of digital materials.

The authenticity of digital materials is threatened whenever they are transmitted across space (i.e., when sent to an addressee or between systems or applications) or time (i.e., either when they are in storage, or when the hardware or software used to store, process or communicate them is updated or replaced). Because the acts of setting aside digital materials for future action or reference and of retrieving them inevitably entail moving them across significant technological boundaries (from display to storage subsystems and vice versa), the inference of the authenticity of digital materials must be further supported by evidence that they have been maintained using technologies and administrative procedures that either guarantee their continuing identity and integrity or at least minimize risks of change from the time the records were first set aside to the point at which they are subsequently accessed.

Technology-independent Authentication

Presumption of Authenticity. A presumption of authenticity is an inference that is drawn from known facts about the manner in which a document has been created and maintained. Adoption and consistent application of the recommendations presented in this document provide the best evidence to support such a presumption. The recommendations are cumulative: the higher the number of satisfied recommendations and the greater the degree to which an individual recommendation has been satisfied, the stronger the presumption of authenticity.

Successful implementation of the recommendations presented in this document is predicated on establishing and continuously applying effective administrative policies and procedures.[18] Ideally, you should strive to implement authentication techniques supported by administrative policies and procedures that are as technology-independent and/or neutral as possible.

[17] Defined as: The properties or attributes conveying the identity of a digital object that is to be kept as a record. (See also Recommendation 3.)

[18] See Appendix 19, "A Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records."

Technology-dependent Authentication

Technology-dependent authentication techniques, such as cryptography, are used to provide a technological mechanism to guarantee the authenticity of digital materials. One such cryptographic technique is the digital signature, which can be used when transmitting documents between persons, systems or applications to declare their authenticity at a certain point in time. Such technologies have been given legal or regulatory value by some bodies, like the European Commission and the Securities and Exchange Commission.

Caution! Digital signatures are subject to obsolescence themselves and, by virtue of their purpose and inherent functionality, cannot be migrated to new or updated software applications together with the documents to which they are attached. In fact, the life of digital signatures and other authentication technologies may be much shorter than the length of time that even a temporary document not requiring migration may need to be maintained, because authentication technology is changing rapidly. Unless or until further development of digital signature technology enables such encrypted authentication information to be preserved over time with the document, you should, when you receive a document with an attached digital signature, detach the signature whenever possible and add information to the integrity metadata to indicate that the document had an attached digital signature when received and that the signature was verified, detached and deleted.

7. Protect digital materials from unauthorized action.

The accuracy and authenticity of digital materials cannot be presumed if there is any opportunity for modifying them without leaving a trace. You need to be able to demonstrate that it was impossible for anyone to tamper with or manipulate your digital materials without that person being identified. Security includes restricting physical access to places where computers are kept, as well as restricting access to the digital materials on the computers themselves. The latter can be accomplished through various means, including the use of passwords and/or biometric authentication to log on to the system.

It is also important to set up a structure of access permissions (also called access privileges—see discussion of *integrity metadata* in Recommendation 4) for all users of the system. For example, some users may only be able to read materials, while others may have permission to modify them. In any case, it should be impossible to modify any record once it has been filed according to the *classification scheme* or filing plan (see Recommendations 3 and 5), and only the person who has been given responsibility for recordkeeping and maintenance should be able to transfer or delete materials from the system. In addition, the system should maintain an audit trail to track access to the materials to control the administration and use of access privileges.

This recommendation may appear to be a tall order for individuals who may be working out of their homes, or even for those working in very small offices or communities of practice. But it is important to remember that if you cannot demonstrate that it was impossible for anyone to tamper with and manipulate your digital materials without being identified, your assertion that your records are de facto accurate and authentic becomes irrelevant. In this regard, it might be useful to keep copies of at least the most important digital materials offline and to establish some routine by which materials stored offline are randomly compared with their counterparts online on a periodic basis.

8. Protect digital materials from accidental loss and corruption.

Computers are not foolproof, and any of a number of factors can cause corruption or other accidental loss of records or data. The best way to ensure against accidental loss or corruption is to make backup copies regularly and often. If you store such copies off-site, additional protection is obtained against fire or theft of equipment. Many backup techniques, software packages, and services are available, including ones that automatically create the backup materials and then transmit them to a secure off-site location.

a. *Develop a rigorous policy or routine that ensures your system is backed up daily*. Your system is only as good as its last backup, so you need to make sure it is backed up often, at least once daily, using proven methods that will ensure that if something goes wrong, you and/or your business will be able to recover quickly. Such regular backups should be destroyed on a rotational basis according to a strategy or schedule that is most appropriate for your requirements, since they do not contain records but only exist for recovery of the system if it fails. Note that we are talking here about a comprehensive *system backup*, which includes the operating system, the software applications and all the digital materials in your system. If, in addition to a system backup, you need to have a security copy of your digital materials in case your computer is stolen or some of your records become corrupted, then you should backup those materials only on another computer, an external hard drive or other portable digital media and store these security copies in an off-site location away from the computer with the "original" copies.
b. *Choose and install the best backup technology for your situation*. Study the technology and services available, and choose what works best for your particular situation. Many different systems are available, ranging from those covering one-person operations to those able to back up very large systems. The backup system needs to include an audit trail, in case the system fails between backups and you need to recover the records or other digital materials created during the time for which there is no backup.

9. Take steps against hardware and software obsolescence.

The speed with which hardware and software become obsolete poses severe challenges to the maintenance and long-term preservation of digital material. One strategy to address this problem is to eliminate dependence on hardware by transferring hardware functionalities to software (i.e., use a software application to simulate the actions of a piece of hardware). This provides a more stable way to retain the function when the hardware becomes obsolete.

The rapidly changing technology environment means that both individuals and offices should regularly upgrade their digital systems as well as all the records within these systems and those that have been moved to another storage medium, such as CD, DVD or tape. In other words, when parts of the technological environment in which you are working begin to become obsolete, they should be upgraded to the most advanced technology available according to your particular requirements and constraints, and all digital materials inside and outside the system should be migrated to the new technology. When replacing hardware, it is important for the replacement hardware to have capabilities at least equal to the hardware it is replacing. For example, a new monitor needs to display a graphic record in a way that retains the documentary form of the original record. Planning for regular technology upgrades on a rotational basis will help ensure that your technology does not become out of date and also help prevent large and unexpected technology expenses.

Sometimes digital records produced by or maintained in systems that are becoming obsolete need to be retained for a long time, but they are not expected to be accessed often. If such records are textual records and need to be read sequentially rather than randomly, you could convert them from their digital form to computer output microfilm. This will protect them from accidental loss or corruption better than any other measure. Another good protective measure is duplication—creating a second copy of groups of vital records and keeping it on another computer, on a second hard drive, on DVD, with another office or individual or in remote storage. When digital records or other entities are removed from a live system, for storage on magnetic or optical media outside the system, for example, it is essential that documentation about the system and about the digital materials (for example, the records' metadata) is also removed and kept with them. For more detailed information about the types of documentation in question here, see Recommendation 1, subsections D, E and F.

10. Consider issues surrounding long-term preservation.

Although the focus of this document has been on the creation and maintenance of all kinds of digital materials while they are needed on a regular basis by their creators, it is important to consider how best to preserve important digital materials for the long term. Typically, only a small percentage of materials need to be preserved for the long term, but the ability to provide ongoing, long-term care for materials, especially digital materials, is often beyond the capability or interest of individuals and small organizations. There are real costs—both financial and human—in retaining materials for the long term, but such preservation efforts are essential for establishing and maintaining our cultural heritage, for accountability purposes and for informing managerial decision-making.

To begin this process, you should identify someone who will take charge of your digital materials once they are no longer needed for regular personal or professional purposes. This person would take the role of *trusted custodian*.[19] A trusted custodian is a professional—or a collection of professionals, as in an archives or a community historical society—who is educated in recordkeeping and preservation, and who ideally has no stake in the content of the records and no interest in allowing others to manipulate or destroy the records.

In the case of small organizations or offices, this person could be the one responsible for keeping the records and organizing and storing them during their active use. In the case of individuals who manage their own recordkeeping, the person fulfilling the preservation function may be an archivist or a librarian in a documentation centre, or simply themselves. In either case, a preservation strategy should be established as soon as possible, because digital materials that have not been targeted for preservation early and taken care of in a proactive way will not be preserved. Close adherence to these guidelines will therefore facilitate long-term preservation.

[19] Defined as: A preserver who can demonstrate that it has no reason to alter the preserved records or allow others to alter them and is capable of implementing all of the requirements for the authentic preservation of records.

Conclusion

This document has outlined a series of activities for individuals and small organizations to carry out to create and maintain digital materials that can be presumed to be authentic, accurate and reliable. For individuals the burden may seem great, but the alternative—loss of records or the emergence of corrupt and unverifiable data—would be an even greater problem in the long run. Small organizations will benefit by making a clear designation of the individual or individuals responsible for overseeing the maintenance of the organization's digital records. Bear in mind, however, that not all recommendations presented in this document need to be implemented in each circumstance; you should be able to select and adopt the measures that address your particular problems in the specific context in which you operate. There may also be cases in which additional measures are necessary because of legislative or regulatory requirements specific to your field, or because of the characteristics of the activity and hence of the records that it produces. In such cases, consultation with experts may be required. Among such experts are the archivists of city, provincial, state or national archives, as well as local archival associations. Individuals, offices and small organizations should not hesitate to contact such experts for advice on any issues relating to the creation and maintenance of their digital materials.

Finally, this set of guidelines is but one of the documents issued by the InterPARES Project, an international research project studying the long-term preservation of authentic digital records. Additional material that will support the understanding of the nature of digital records and the development of methods for their reliable creation and accurate and authentic maintenance and preservation can be found on the InterPARES Web site at www.interpares.org.

APPENDIX 21

PRESERVER GUIDELINES
Preserving Digital Records: Guidelines for Organizations

by

Yvette Hackett
Library and Archives Canada

PRESERVER GUIDELINES
Preserving Digital Records: Guidelines for Organizations[1]

Introduction

These guidelines have been developed to provide concrete advice to various groups that are responsible for the long-term preservation of digital records. They are not intended to be comprehensive but to highlight a number of areas that are particularly important to the preservation of authentic digital records and which experience has shown to be often overlooked in the rush to accept digital records into archival repositories.

As is widely recognized, digital records must be carefully managed throughout their entire existence to ensure that they are accessible and readable over time with their form, content and relationships intact to the extent necessary for their continuing trustworthiness as records. It is also widely recognized that management of digital records must proceed from a comprehensive understanding of all phases or stages of records' existence, from the time they are generated, through their maintenance by their creator, and during their appraisal, disposition and long-term preservation as authentic memorials of the actions and matters of which they are a part. From the perspective of long-term preservation, all the activities to manage records throughout their existence are linked, as in a chain, and interdependent. If a link in the chain fails, the chain cannot do its job. If certain activities and actions are not undertaken on records, their integrity (that is, their reliability and authenticity) and preservation are imperilled.

These guidelines focus on the preservation link in the chain of preservation and are organized according to the sequence of preservation activities presented in the InterPARES Chain of Preservation (COP) model,[2] which charts the many sequential steps in the creation, maintenance and preservation of authentic records. The alphanumeric number in parentheses following each section title in these Guidelines is a cross reference to the applicable preservation activity presented in the COP model.

The guidelines have been tailored to address the preservation needs of organizations or programs whose records must be retained and consulted for long periods and those of archival institutions that take on the responsibility for the long-term preservation of the records of others and for their continuing accessibility to the public they serve. In both these cases, human and financial resources as well as in-house technical expertise are frequently limited.

Institutions, organizations and programs with preservation responsibilities should also consult the *Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records* (a.k.a., Policy Framework)[3] developed by the InterPARES 2 Policy Cross-domain, which complement these Guidelines. Many of the recommendations of these Guidelines may also be applicable to the preservation of digital objects other than records, such as documents, publications or data.

---

[1] These Guidelines have also been issued in an illustrated booklet form that is freely available at http://www.interpares.org/display_file.cfm?doc=ip2(pub)preserver_guidelines_booklet.pdf.

[2] Available at http://www.interpares.org/ip2/ip2_models.cfm.

[3] Available at http://www.interpares.org/public_documents/ip2(pub)policy_framework_document.pdf.

1. Manage Chain of Preservation

This aspect involves determining framework requirements, and designing, implementing and maintaining a chain of preservation framework. A *Chain of Preservation Framework* includes all the elements of policy, strategy, methodology and so on.

1.1. Establish scope and objectives

Preservers must define the scope and objectives of their digital preservation program. In the arts, for example, they may wish to preserve the recording of the performance(s) of a work, or they may choose to undertake the more complex preservation of the components of a work of art that support its reproduction or re-performance. In the sciences, preservers may wish to preserve only the final report of the results of an experiment, or hold raw data, normalized data and/or aggregated data to document the methodology used and the result obtained, as well as to ensure the availability of the data for future uses. Preservers should also consider who the eventual users of the archives will be. Technically sophisticated users generally require less assistance in accessing even technologically complex digital materials, while the general public might require extremely user-friendly access mechanisms and materials transformed into a few simple, but widely available, formats. The scope of the preservation program will help define which preservation strategies (see Section 4 and Appendix 21c, Section B) a preserver might need to support.

In defining the digital preservation program, preservers should build on previous efforts. To develop appropriate policies and strategies, preservers should consult the InterPARES 2 Policy Framework for guidance applicable at organizational, sectoral, national, international and supranational levels. For the functions of the preservation program, preservers should consult the ISO Open Archival Information System (OAIS) standard[4] and should follow the InterPARES 2 Chain of Preservation model for an adaptation of the OAIS standard specifically intended for digital records. Plans should also reflect the *Trustworthy Repositories Audit & Certification: Criteria and Checklist*, a revised and expanded version of the *Audit Checklist for Certifying Digital Repositories* originally developed by the NARA/RLG Digital Repository Task Force.[5]

1.2. Acquire resources

Digital preservation requires substantial resources in funding, technological capabilities and expertise. An organization responsible for digital preservation has several options, including: a) acquire new resources, b) reallocate existing resources and/or c) leverage other resources.

Regardless of the option(s) chosen, a fundamental requirement is that resources must be sustainable. One-time resources, such as grants, may be appropriate for specific finite tasks, such as establishing the preservation program or processing a given body of records, but a reliable source of sustained resources is a *sine qua non* for any preservation program.

Acquiring new financial resources will require a sound plan for the program and a matching communications plan to convince funding sources and stakeholders that preservers are likely to consult that the program should be funded. A viable strategy for a new program may be to start small and plan on short-term successes to convince funding sources to incrementally increase

[4] International Organization for Standardization, ISO 14721: 2003 - Space data and information transfer systems—Open archival information system—Reference model.
[5] See Online Computer Library Center, Center for Research Libraries (2007), "Trustworthy Repositories Audit & Certification: Criteria and Checklist," v. 1.0, February 2007. Available at http://www.crl.edu/PDF/trac.pdf.

resources for the program. An incremental strategy should evaluate whether funding sources are more likely to be influenced by short-term success in basic program objectives or in areas of more particular concern to the funding sources and stakeholders. For example, funders and stakeholders may be more swayed by demonstrations of technological capabilities than by a sound and comprehensive plan for appraising digital records.

For most organizations, reallocating resources to digital preservation is likely to entail painful choices. As with seeking new funds, an incremental approach may be best. Furthermore, ongoing adjustments can be made to the plan, based on the experience gained during each phase of implementation. If the digital preservation program is to be established in a larger institution, it would be helpful to address digital preservation as part of the overall strategic plan rather than as a special initiative.

Even when a preserver successfully acquires new resources or is able to reallocate existing resources to digital preservation, it is unlikely it will have sufficient resources to address all the challenges. Therefore, preservers should capitalize on opportunities for leveraging outside resources. There are a variety of paths for doing this. For example, rather than trying to hire technical experts on a permanent basis or training staff in all requisite technical knowledge and skills, preservers might engage outside experts on a consultative or task basis. They should not exclude options to contract for both basic and ad hoc services. On a basic level, preservers should evaluate the possibility of using a computer service provider rather than acquiring a dedicated preservation system. Ad hoc options include engaging specialized companies for tasks such as re-copying from obsolete digital media or converting rare formats. Another option is to participate—actively or passively—in open-source communities developing technologies needed for digital preservation (e.g., FEDORA,[6] Global Registry of Digital Formats[7]).

Finally, preservers in an organization lacking the required resources to support a digital preservation program should investigate the possibility of establishing collaborative partnerships or consortia to develop and finance a program that meets a minimum acceptable standard.

1.3. Focus on digital records

Preservers must ensure that digital preservation resources are primarily deployed to protect authoritative copies[8] of digital records, rather than to preserve digitized copies of surviving analogue records. The rationale for this is that most analogue records will survive without digitization, whereas digital records will be lost without a digital preservation program.

1.4. Offer advice

Because the chain of preservation of digital records begins at creation, preservers should provide advice on digital records creation and maintenance. Depending on the mandate of the preserver, this may be quite specifically targeted to, for example, employees in the preserver's organization or, as in the case of national archives, other government institutions. In other cases, the advice may be disseminated widely to special interest groups or to the general public, with the purpose of reaching the person(s) or organization(s) whose records fall under the mandate of the preserver.

---

[6] See http://hul.harvard.edu/formatregistry/.
[7] See http://www.fedora.info/.
[8] Authoritative copy is defined as "The instantiation of a record that is considered by the creator to be its official record and is usually subject to procedural controls that are not required for other instantiations" (InterPARES 2 Terminology Database. Available at http://www.interpares.org/ip2/ip2_terminology_db.cfm).

1.5. Set a good example

Preservers must establish, within their own organization, a record-making and a recordkeeping environment such that their own control records produced in the course of their preservation function will be created and maintained in a way that satisfies the InterPARES 1 *Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records*.[9] Not only is this an essential requirement for any organization undertaking long-term preservation, but the development of this type of in-house environment will also provide:

- hands-on training to archivists in the technologies they are championing to records creators;
- an invaluable "user's eye view" of actual recordkeeping solutions and how they really work in a day-to-day operational environment;
- a testbed where upgrades and innovations can be introduced and evaluated; and
- a working prototype that can be used in demonstrations.

1.6. Develop procedures

Preservers must establish controls over records transfer, maintenance and reproduction, including the procedures and system(s) used to transfer records to their own organization or program within the organization; maintain them; and reproduce them in a way that satisfies the InterPARES 1 *Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records*.[10] These procedures must embody adequate and effective controls to guarantee the records' identity[11] and integrity,[12] and specifically that:

- unbroken custody of the records is maintained;
- security and control procedures are implemented and monitored;
- the content of the records and the required annotations and elements of documentary form remain unchanged after reproduction.

1.7. Implement maintenance strategies

Although much attention is paid to the development of complex long-term preservation strategies, such strategies are inapplicable if the records for which they are to be used are not properly maintained and protected in the recordkeeping and/or record preservation systems that contain them. A complete version of the eight primary maintenance strategies is available in Appendix 21c, Section A. Briefly, they include:

A1. Clear allocation of responsibilities
A2. Provision of appropriate technical infrastructure
A3. Implementation of a plan for system maintenance, support and replacement

---

[9] See Authenticity Task Force (2002), "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 204-219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf. See Appendix 21a for an abridged version.
[10] Ibid. See Appendix 21b for an abridged version.
[11] Identity is defined as "The whole of the characteristics of a document or a record that uniquely identify it and distinguish it from any other document or record. With integrity, a component of authenticity" (InterPARES 2 Terminology Database, op. cit.).
[12] Integrity is defined as "The quality of being complete and unaltered in all essential respects. With identity, a component of authenticity" (Ibid.).

A4. Implementation of a plan for the transfer of records to new storage media on a regular basis
A5. Adherence to appropriate storage and handling conditions for storage media
A6. Redundancy and regular backup of the digital objects
A7. Establishment of system security
A8. Disaster planning

2. Appraise Records for Permanent Preservation (A4.2)

In cases where, as recommended in the InterPARES 2 Chain of Preservation model, retention scheduling is employed, decisions on the disposition of records will regularly be made as part of the management of a recordkeeping system. In some cases, appraisals may be conducted when it is determined that records in a longstanding system need to reach a disposition. Eight important aspects of the appraisal process are discussed below.

2.1. Appraise early

Given the technical difficulties involved in the preservation of digital records, the identification of what records need to be preserved for the long term should be carried out at the earliest possible opportunity. Performing appraisal, establishing transfer methods and even identifying potential preservation strategies with the records creator will improve the likelihood of success. This process may also provide the preserver with an opportunity to offer records creation and maintenance advice (see Section 1.4).

Professional preservers, such as archivists, are frequently encouraged to participate in the actual design of computer applications being developed by organizations with which they have a donor-preserver relationship. This approach will help integrate appropriate recordkeeping and preservation practices. Preservers who have joined system design teams have learned that it is an enormously time-consuming practice that requires a far more detailed understanding of the organization's internal workflows and procedures than an archivist normally acquires during an appraisal. Furthermore, system specifications are rarely an accurate depiction of the application that will eventually be implemented. An appraisal will still have to be conducted once the system is operational and is meeting organizational requirements. It may be more reasonable for archivists to contribute to system design as part of the advice function discussed in Section 1.4. Sharing high level strategies, principles and guidelines developed by the archival profession may prove to be a more realistic goal.[13]

2.2. Locate multiple owners

In cases where the intellectual components of a digital object have multiple owners, these owners must be identified during the appraisal process to assess the ramifications of this situation

[13] Many aspects relating to the creation of effective digital preservation programs have been studied in recent years. Among the Web sites containing useful information or examples are: the InterPARES Project at http://www.interpares.org; Model Requirements for the Management of Electronic Records (MoReq) at http://www.cornwell.co.uk/edrm/moreq.asp; the Metadata Encoding and Transmission Standard (METS) at http://www.loc.gov/standards/mets/; the Electronic Records from Office Systems (EROS) at the National Archives of the United Kingdom at http://www.nationalarchives.gov.uk/electronicrecords/advice/guidelines.htm; and the Australian DIRKS (Designing and Implementing Recordkeeping Systems) manual at http://www.records.nsw.gov.au/recordkeeping/dirks-manual_4226.asp.

for long-term preservation. This can occur, for example, where institutions at various levels of government contribute, and share access to, data resources. Another example is illustrated by Web sites that access and use resources located outside their span of control. Although access agreements are frequently negotiated in these circumstances, they rarely include provisions for long-term preservation of all significant digital components.

2.3. Assess authenticity

The assessment of authenticity has always formed part of the traditional archival appraisal process. In the first instance, it has relied on confirming the existence of an unbroken chain of custody from the time of the records' creation to their transfer to the archival entity responsible for their long-term preservation. Periods when records were not subject to some form of protective measures by the records creator or by a successor institution with a vested interest in maintaining the accuracy and completeness of the records can cast significant doubt on the authenticity of the records.

The assessment of authenticity has also depended on the archivist's knowledge of recordkeeping practices, both historically and in relation to the record types and administrative procedures of a specific creator. The general framework for this assessment was originally codified in diplomatics.[14] A third, less frequently used method to confirm the identity and integrity of records is based on comparison. Records within a fonds are compared to copies forwarded to and held by external sources in the normal course of the creator's business.

Records created and maintained using digital technology present additional difficulties, and archivists have not yet developed standard practices to assess authenticity in this environment. Issues revolve around the fact that digital objects are easily duplicated, distributed, re-named, re-formatted or converted, as well as to the ease with which they can be falsified without leaving a trace. The following examples illustrate the extent of the loss to archivists, historians, lawyers and others who require authentic records in their work:

- The physical support on which digital documents are stored has largely lost its significance in confirming the date of a record or its place of manufacture. Anyone with access to functioning, obsolete equipment and storage media has the capability to copy digital files to, for example, 9-track tape or 5-1/4" diskettes.
- The date stamp on any digital file can be modified by adjusting the system clock.
- Few institutions understood what their employees would do once entrusted with word processing software. Standard document forms, such as memos and correspondence on letterhead, disappeared under the onslaught of new, individualized record forms, which rapidly included personalized colour, graphics and even sound effects, as well as the attribution of new meaning to capitalization, colour and the development of emoticons. The degree of erosion of standard records creation practices varied enormously across types and sizes of corporate and government organizations.
- The introduction of e-mail networks allowed records to travel by many new routes among staff, rather than according to the well-established distribution routes of traditional office procedures.
- The severe reductions in records management personnel in most organizations, fuelled by an assumption that digital objects somehow did not need to be managed, played havoc

[14] See discussion of diplomatics in Luciana Duranti and Kenneth Thibodeau (2006), "The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES," *Archival Science* 6(1): 15-21.

with the holdings of the Records Office, which largely stopped receiving records created and transmitted in digital form.

When appraising records created in a digital environment, the assessment of the authenticity of records must become a more overt, visible process performed and documented by the preserver. Unbroken chain of custody, knowledge of recordkeeping practices, and verification may still offer some assurances of authenticity. To these must now be added the verification of compliance with each of the benchmark requirements for authenticity listed in Section 2.4.

2.4. Document the assessment of authenticity

The appraisal report should document the controls put in place by the creator to guarantee the identity and integrity of the records and thus the presumption of their authenticity. These controls include each of the benchmark requirements supporting the presumption of authenticity.[15] Briefly, these include:

A.1 Expression of Record Attributes and Linkage to Record (e.g., identity and integrity metadata)
A.2 Access Privileges
A.3 Protective Procedures against Loss and Corruption of Records
A.4 Protective Procedures against Media Deterioration and Technological Change
A.5 Establishment of Documentary Forms
A.6 Authentication of Records
A.7 Identification of Authoritative Record
A.8 Removal and Transfer of Relevant Documentation

2.5. Monitor records identified for long-term preservation

Once the appraisal is completed, the records identified for preservation must be monitored at regular intervals until such time as they will be transferred to the preserver. Monitoring involves confirming with the records creator that nothing has changed with regard to how classes of records identified for transfer are being created or maintained or, if changes have occurred, that they have not affected the nature and attributes of the records, their value, their authenticity or the feasibility of their preservation.

Many changes within an organization can affect the ongoing survival of digital records. The possibility that records will be destroyed in an instant is much higher than for traditional records. This danger is somewhat offset by the tendency to duplicate material in an uncontrolled fashion. Unfortunately, if the production of copies is uncontrolled, it is unlikely that anyone will realize when the last copy of a record is destroyed.

The simplest scenario may involve a system upgrade either to the hardware or to the software, which will affect the archives' ability to accept the records. An upgrade could also result in even minor system re-design that could remove the ability to separate temporary records from those that must be removed for transfer to the preserver.

A second scenario can involve changes in an organization's mandate or functions. This can easily lead to changes in how computer applications are used, and the nature and amount of data that they contain. People responsible for system re-design may not be aware of the requirement for transfer of the existing records to the designated preserver before the system can be modified.

[15] See Appendix 21a.

Without intervention, even documentation about the original application and backup tapes will move inexorably toward a scheduled destruction date.

Finally, the widespread collapse of proper records management practices in most organizations means that records are poorly identified and incorrectly stored in unsecured locations. Managers, and even records managers, may not understand the details of the technical infrastructure, while IT staff may be unfamiliar with either the history of an organization or the relative importance of older records in various data stores. Hard drives may be wiped, user accounts and all the files they contain may be deleted, tapes and discs may be recycled or destroyed, and obsolete playback technology may be disposed of to meet day-to-day operational requirements of speed and efficiency, with no understanding of the impact of such actions on an organization's records or on pre-existing transfer agreements designed to ensure their long-term preservation.

2.6. Update appraisals

Appraisals also need to be updated at regular intervals, though less frequently than records identified for transfer need to be monitored. Information gathered during a monitoring visit may provide the first indication that a new appraisal is required. Change within organizations and within their record-making and recordkeeping systems is inevitable. Organizational mandates and responsibilities may change, as well as the way those responsibilities are carried out, and data accumulated in existing systems may be put to new uses, which might increase their long-term value. At the simplest level, systems that did not initially contain records may be upgraded to do so. This is particularly true during this period of "hybrid" recordkeeping systems, where paper-based record systems continue to co-exist with the early stages of digital information, document or record systems.

2.7. Identify all digital components[16]

Paper records kept in traditional recordkeeping systems generally offer a tightly-wrapped package, where the content of the record is firmly attached to its paper support and the record itself is contextually filed with the related records. This seamless system began to break down with the introduction of technology when, for example, photographic negatives had to be processed to produce prints and moving images resulted from multiple layers of sound and images, combined and re-combined to produce the final composite print that is screened in cinemas or broadcast on television.

Digital technology has further dismantled the record into a series of components. To successfully extract digital records from the system in which they were created, or even from a secondary maintenance system, the preserver must ensure that all essential digital components are identified and that implicit relationships are made explicit in the metadata before the whole construct is transferred. One of the most common examples of a digital component is the library of fonts, any number of which can be selected by the creator to be used in the presentation of a word-processed document. In Windows, these are stored in '.dll' (or dynamic link library) files. For the preserver to be able to reproduce this record to reflect the creator's original intentions, both the digital component containing the text and the digital component containing the font

[16] A digital component is defined as "A digital object that is part of one or more digital documents, and the metadata necessary to order, structure or manifest its content and form, requiring a given preservation action" (InterPARES 2 Terminology Database, op. cit.).

must have been preserved, as well as the link between them established in such a way that the software attempting to display the content of the text file can find the appropriate font library.[17]

2.8. Determine the feasibility of preservation

Although not part of the assessment of the value of the records, the appraisal process must be completed by a careful investigation of the technical preservation requirements for preservation. Different preservation strategies (see Appendix 21c, Section B) can vary widely in cost and can produce very different results. A textual record stripped of all its formatting may be acceptable in a situation where the preserver is interested in carrying forward only the content of the record. However, where meaning is conveyed by the documentary form and the display characteristics of the record, a more complex preservation solution will be required.

A determination of the feasibility of preservation is essential if the preserving body is to clearly understand the cost of the acquisition and preservation to which it is committing. This is not a new activity; it is simply the extension to the digital realm of the identification of the resources needed to preserve, for example, paper records that are mouldy or moving image reels that are badly shrunken. The current state of digital preservation does mean, however, that preservation costs must be viewed as recurrent. Re-copying holdings from one physical carrier to another will be required as often as the selected format becomes obsolete. Conversion of file formats will be required when logical obsolescence threatens to make the content unreadable. In addition, the digital records considered for long-term preservation may require measures far too complex for the technological environment and the knowledge resources of the preserving organization, and this might imply a postponement of the transfer.

3. Acquire Selected Records for Permanent Preservation (A4.3)

The activity of the preserver acquiring selected records, and all the activities of preservation that follow from that, have as their goal the continued authenticity and accessibility of those records that are selected for continuing preservation. This movement of records from the creator's (or legitimate successor's) custody to the preserver's custody is a critical juncture in the chain of preservation and must be done with great care to ensure that nothing goes awry in the transfer process.

3.1. Develop shared plan for transfer

A successful transfer from the current custodian of the records (be it original creator or legitimate successor) to the organization or program taking on responsibility for long-term preservation requires a plan agreed upon by both parties. Re-accessing obsolete systems or extracting inactive records from operational systems will definitely involve human resource costs for copying time and, potentially, for programming time. Special hardware and software may also be required. The logical and physical (or virtual) formats used for the transfer must be agreeable to both parties. As a general rule, the transfer plan should be developed when the technical feasibility of acquisition and preservation are undertaken. If the two parties cannot

[17] A more detailed description of the "digital component," with additional examples illustrating the concept, is available in Preservation Task Force (2001), "Appendix 6: How to Preserve Authentic Electronic Records," in Duranti, *Long-term Preservation*, op. cit., 293–328. Online reprint available at http://www.interpares.org/book/interpares_book_o_app06.pdf.

agree on a transfer process, the appraisal decision may have to be re-visited. Again, in this period of hybrid recordkeeping, paper and microfilm-based options may still exist. Alternatively, the preserver might encourage the records creator to adopt upgrades to the record system that will allow for easier regular transfers.

3.2. Enforce standardized procedures

The controls over the transfer of digital records from the creator's to the preserver's custody must include:

- establishing, implementing, and monitoring procedures for registering the records transfer;
- verifying the authority for transfer;
- examining the records to determine whether they correspond to the records that are designated for transfer; and
- accessioning the records.

As part of the transfer process, the authenticity of the creator's records, which was assessed as part of the appraisal process, should be verified. This includes verifying that the metadata relating to the records' identity and integrity have been transferred together with the related records and are linked to them, and that the records are accompanied by any relevant documentation of the technical and administrative environment in which they were created and maintained.

3.3. Keep the oldest available logical format

The logical format[18] in which the records were originally created, or in which they are held by the creator at the time of transfer, should, whenever feasible, be maintained by the preserver, in addition to any preservation or reference copies generated after the transfer. Should selected preservation strategies, such as a specific conversion path, fail over time, continued custody of the initial logical format will allow the preserver to essentially re-start the preservation process with the most authoritative copy of the records, by applying a different preservation strategy to the records. Over the long periods during which preservers hold records, experience may show that other preservation strategies are more stable over time or can more easily be carried forward over the long-term. Alternately, new methods of preservation may have been developed following the acquisition and initial processing of the records.

3.4. Avoid duplicates

Because of the ease of replication of digital records, the preserver must put in place procedures to ensure that digital records from a specific series are transferred by a specific creator to the preserver only once. Accurate identity information is an important first step in avoiding duplication of effort by the creator and the preserver. Also, if reference copies are provided by the preserver to the creator after the transfer of the records, they should be clearly identified and marked as such to prevent accidental re-transfer.

[18] Logical format is defined as "The organized arrangement of data on electronic media that ensures file and data control structures are recognizable and recoverable by the host computer operating system" (InterPARES 2 Terminology Database, op. cit.). Two common logical formats for files and directories are ISO 9660 for CD-ROMs, and Universal Disk Format (UDF) for DVDs.

3.5. Document all processing

Initial processes applied during and immediately after transfer may or may not be related to preservation per se. Confirming the identity of the transferred material, checking for viruses and confirming completeness of files tend to leave the transferred file unchanged. File conversion, renaming digital objects and encapsulating files are more intrusive activities. In both cases, preservers must document all processing of digital records and the effects of processing while records are in their custody (see Appendix 21b, Requirement B.2). This documentation should include information such as:

- why certain processes were applied to the records;
- what records were processed;
- the date when the process was performed;
- the names of persons performing and documenting the various steps of the process(es);
- the impact of the process performed on the records' form, content, accessibility and use; and
- the description of any damage, loss or other problems encountered as a result of the processing, including any effect on the elements expressing the records' identity and integrity.

Should the preserver produce copies of the acquired records, it is important to remember that, as discussed in Section 1.5, these copies should be produced in an environment that satisfies the relevant requirements[19] from the InterPARES 1 Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records.

4. Preserve Accessioned Records (A4.4)

The designated records preserver is the entity responsible for taking physical and legal custody of, and preserving (i.e., protecting and ensuring continuous access to), a creator's records. Be it an outside organization or an in-house unit, the role of the designated preserver should be that of a trusted custodian[20] for a creator's records. The authentic copies of the creator's records are kept by the trusted custodian in a *trusted preservation system* (see Appendix 21c), which should include in its design a description and a retrieval system. This trusted preservation system must also have in place rules and procedures for the ongoing production of authentic copies as the existing system becomes obsolete and the technology is upgraded.

4.1. Describe the records

The information about the records and their contexts collected during the appraisal and processing stages should form part of the archival description of the fonds or series in which the records belong (see Appendix 21b, Requirement B.3). This should also include information about intellectual property rights or privacy concerns.

The archival description of the fonds or series containing the digital records should include—in addition to information about the records' juridical-administrative, provenancial, procedural,

[19] Requirement A.5 (Establishment of Documentary Forms), where the creator establishes the documentary form of the record, would usually not apply to the preserver, except if the original documentary form of the record has been lost and the preserver must specify a substitute to permit access.

[20] A trusted custodian is defined in the InterPARES 2 Terminology Database as "A preserver who can demonstrate that it has no reason to alter the preserved records or allow others to alter them and is capable of implementing all of the requirements for the authentic preservation of records" (InterPARES 2 Terminology Database, op. cit.).

and documentary contexts—information about changes the digital records of the creator have undergone since they were first created. The description should also include an overview of the transfer and preservation processes based on the documentation discussed in Section 3.5 and the explanation of the relationships among digital components discussed in Section 2.7.

4.2. Identify legal ramifications of preservation actions

When a preservation strategy is selected, its legal implications should be reviewed. For example, format conversion out of a proprietary environment could involve the preserver in illegal actions. In the United States, the *Digital Millennium Copyright Act* has made it a criminal offence to produce tools that can circumvent copyright protection measures. Internationally, the World Intellectual Property Organization Copyright Treaty (WIPO WCT) contains provisions that include copyright protection for software as well as digital works and that introduce criminal penalties for infringement, which ranges from unauthorized copying of material placed on a Web site to the removal or alteration of rights management controls from digital works. Most software packages also include some type of similar restrictions, which users must agree to during the installation process.

4.3. Confirm the effectiveness of the selected preservation strategy

As discussed in Section 2.8, there are now a number of preservation strategies available. Ideally, the selected preservation strategy should be tested on the records prior to the formal transfer to the preserver, to ensure that it will perform as expected. Realistically, most preserving organizations or programs can only fund this type of testing on an exceptional basis. Just as traditional conservators carefully test proposed treatments before applying them wholesale to analogue records, digital preservers must be constantly alert to the impact that each preservation process may have on the records and ensure that it is the appropriate choice for preserving authentic records. Flaws in application software and variations in the functionality of versions over time can result in unexpected consequences when applied to a new group of records.

Part of this process includes a constant awareness of the need to track the presence and the performance of all digital components. A change in one component may have unexpected results on a second component, or it may affect how the relationship functions between any two essential components of the record or affect these components' ability to interact. A different relationship that could be affected is that which exists among the members of a related group of records, such as a dossier or series, and the presentation of that aggregate in the correct order (e.g., alphabetical, chronological or hierarchical). If the original order has been lost, corrective measures will have to be taken.

4.4. Maintain proper storage

It is a widely accepted archival preservation principle that maintaining an appropriate and consistent storage environment (temperature and relative humidity) for the material being stored is the most cost-effective contribution to the long-term preservation of records. Manufacturers of magnetic or optical storage media generally offer advice on optimum storage conditions. The environment must be monitored constantly and the readings checked on a regular basis. This recommendation is one of the eight mandatory maintenance strategies outlined in Section 1.7 and discussed in Appendix 21c, Section A.

5. Output Records (A4.5)

As noted earlier, continued accessibility (i.e., use) is an integral part of the archival process. Consequently, providing access to preserved records is an essential component in the chain of preservation. It should be managed by the preserver with the same sense of responsibility and degree of technical and professional competence imparted to records appraisal, acquisition/ transfer, description and storage.

5.1. Explain how the reference copies were made

The relationship between the records acquired from the creator and any copies produced by the preserver must be clearly described and readily accessible to users (see Appendix 21b, Requirement B.2.b). This should also include documenting how the reproduction process control measures that are in place were established and implemented and how they are monitored to ensure that the content of the reproduced records is not changed in the course of reproduction. Copies of records in the preserver's preservation system may not be designated authentic if the preserver has made them for purposes other than preservation; for example, a copy from which personal identifiers are removed may be made for access purposes.

Documenting the records reproduction process and its effects is an essential means of demonstrating that the reproduction process is transparent (i.e., free from pretence or deceit). Such transparency is necessary to the effective fulfillment of the preserver's role as a trusted custodian of the records. It also provides users of the records with a critical tool for assessing and interpreting the records by demonstrating the continuing authenticity of the records and by providing a complete history of the records, of which the history of reproduction is an essential part.

5.2. Explain the technical requirements for access

As mentioned in Section 1.1, different preservers provide reference services to different types of users. This will affect the reference formats and mechanisms adopted by the preserving organization or program, with simpler methods required for members of the general public who may not even own a computer or who may own a fairly simple machine with a few standard pieces of software. To meet the needs of these users, the preserver may have to undertake additional processing or create specialized tools to assist the researchers. More technologically adept users, such as statisticians doing data analysis or forensic accountants conducting fraud investigations, are more likely to apply their own software tools to copies of the records.

Conclusion

This document has outlined a series of guidelines for institutions, organizations and programs with preservation responsibilities for digital records that can be presumed to be authentic and accurate while in the custody of the preserver. For individual preservers and small preservation organizations, the burden may seem great, but the alternative—loss of records or the emergence of corrupt and inauthentic records—would be an even greater problem in the long run. Small organizations will benefit by making a clear designation of the individual or individuals responsible for overseeing the preservation of the organization's digital records. Bear in mind, however, that not all recommendations presented in this document need to be implemented in each circumstance; each preserver should be able to select and adopt the measures that address its particular problems in the specific context in which it operates. There may also be cases in which additional measures are necessary because of legislative or regulatory requirements specific to the preserver's administrative or cultural jurisdiction. In such cases, consultation with legal experts may be required. Individuals, offices and small organizations responsible for preservation should not hesitate to contact such experts for advice on any issues relating to the preservation of the digital records in their custody and under their control.

APPENDIX 21A

Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records

Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records[1]

Preamble

The benchmark requirements are the conditions that serve as a basis for the preserver's assessment of the authenticity of the creator's electronic records. Satisfaction of these benchmark requirements will enable the preserver to infer a record's authenticity on the basis of the manner in which the records have been created, handled and maintained by the creator.

Within the benchmark requirements, Requirement A.1 identifies the core information about an electronic record—the immediate context of its creation and the manner in which it has been handled and maintained—that establishes the record's identity and lays a foundation for demonstrating its integrity. Requirements A.2–A.8 identify the kinds of procedural controls over the record's creation, handling and maintenance that support a presumption of the record's integrity.

Benchmark Requirements (Requirement Set A)

To support a presumption of authenticity the preserver must obtain evidence that:

REQUIREMENT A.1: Expression of Record Attributes and Linkage to Record	the value of the following attributes are explicitly expressed and inextricably linked to every record. These attributes can be distinguished into categories, the first concerning the identity of records, and the second concerning the integrity of records.

A.1.a Identity of the record:

A.1.a.i Names of the persons concurring in the formation of the record, that is:
- name of author[2]
- name of writer[3] (if different from the author)
- name of originator[4] (if different from name of author or writer)
- name of addressee[5]

A.1.a.ii Name of action or matter

A.1.a.iii Date(s) of creation and transmission, that is:
- chronological date[6]

[1] Excerpted from: Authenticity Task Force (2002), "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 204-219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf.
[2] The name of the physical or juridical person having the authority and capacity to issue the record or in whose name or by whose command the record has been issued.
[3] The name of the physical or juridical person having the authority and capacity to articulate the content of the record.
[4] The name of the physical or juridical person assigned the electronic address in which the record has been generated and/or sent.
[5] The name of the physical or juridical person(s) to whom the record is directed or for whom the record is intended.
[6] The date, and possibly the time, of compilation of a record included in the record by the author or the electronic system on the author's behalf.

- received date[7]
- archival date[8]
- transmission date(s)[9]

A.1.a.iv Expression of archival bond[10] (e.g., classification code, file identifier)

A.1.a.v Indication of attachments

A.1.b Integrity of the record:

A.1.b.i Name of handling office[11]

A.1.b.ii Name of office of primary responsibility[12]
(if different from handling office)

A.1.b.iii Indication of types of annotations added to the record[13]

A.1.b.iv Indication of technical modifications;[14]

REQUIREMENT A.2: Access Privileges	the creator has defined and effectively implemented access privileges concerning the creation, modification, annotation, relocation, and destruction of records;

REQUIREMENT A.3: Protective Procedures: Loss and Corruption of Records	the creator has established and effectively implemented procedures to prevent, discover, and correct loss or corruption of records;

REQUIREMENT A.4: Protective Procedures: Media and Technology	the creator has established and effectively implemented procedures to guarantee the continuing identity and integrity of records against media deterioration and across technological change;

[7] The date, and possibly the time, when a record is received by the addressee.
[8] The date, and possibly the time, when a record is officially incorporated into the creator's records.
[9] The date and time when a record leaves the space in which it was generated.
[10] The archival bond is the relationship that links each record, incrementally, to the previous and subsequent ones and to all those [that] participate in the same activity. It is originary (i.e., it comes into existence when a record is made or received and set aside), necessary (i.e., it exists for every record), and determined (i.e., it is characterized by the purpose of the record).
[11] The office (or officer) formally competent for carrying out the action to which the record relates or for the matter to which the record pertains.
[12] The office (or officer) given the formal competence for maintaining the authoritative record, that is, the record considered by the creator to be its official record.
[13] Annotations are additions made to a record after it has been completed. Therefore, they are not considered elements of the record's documentary form.
[14] Technical modifications are any changes in the digital components of the record as defined by the Preservation Task Force. Such modifications would include any changes in the way any elements of the record are digitally encoded and changes in the methods (software) applied to reproduce the record from the stored digital components; that is, any changes that might raise questions as to whether the reproduced record is the same as it would have been before the technical modification. The indication of modifications might refer to additional documentation external to the record that explains in more detail the nature of those modifications.

REQUIREMENT A.5: Establishment of Documentary Forms	the creator has established the documentary forms of records associated with each procedure either according to the requirements of the juridical system or those of the creator;

REQUIREMENT A.6: Authentication of Records	if authentication is required by the juridical system or the needs of the organization, the creator has established specific rules regarding which records must be authenticated, by whom, and the means of authentication;

REQUIREMENT A.7: Identification of Authoritative Record	if multiple copies of the same record exist, the creator has established procedures that identify which record is authoritative;

REQUIREMENT A.8: Removal and Transfer of Relevant Documentation	if there is a transition of records from active status to semi-active and inactive status, which involves the removal of records from the electronic system, the creator has established and effectively implemented procedures determining what documentation has to be removed and transferred to the preserver along with the records.

Commentary on the Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records

The assessment of the authenticity of the creator's records takes place as part of the appraisal process. That process and the role of the benchmark requirements within it are described in more detail in the "Appraisal Task Force Report." This assessment should be verified when the records are transferred to the preserver's custody.

A.1 Expression of Record Attributes and Linkage to Record

The presumption of a record's authenticity is strengthened by knowledge of certain basic facts about it. The attributes identified in this requirement embody those facts. The requirement that the attributes be expressed explicitly and linked inextricably[15] to the record during its life, and carried forward with it over time and space, reflects the task force's belief that such expression and linkage provide a strong foundation on which to establish a record's identity and demonstrate its integrity. The case studies undertaken as part of the work of the task force revealed very little consistency in the way the attributes that specifically establish the identity of a record are captured and expressed from one electronic system to another. In certain systems, some attributes were explicitly mentioned on the face of the record; in others they could be found in a wide range of metadata linked to the record or they were simply implicit in one or more of the record's contexts. In many cases, certain attributes (e.g., the expression of the archival bond) were not captured at all. The task force's concern is that, in the absence of a precise and explicit

[15] For the purposes of this requirement, *inextricable* means incapable of being disentangled or untied, and *link* means a connecting structure.

statement of the basic facts concerning a record's identity and integrity, it will be necessary for the preserver to acquire enormous, and otherwise unnecessary, quantities of data and documentation simply to establish those facts.

The link between the record and the attributes listed in Requirement A.1 is viewed by the task force as a *conceptual* rather than a *physical* one, and the requirement could be satisfied in different ways, depending on the nature of the electronic system in which the record resides. For example, in electronic records management systems, this requirement is usually met through the creation of a record profile.[16] In other types of systems, the requirement could be fulfilled through a topic map. A topic map expresses the characteristics (i.e., *topics*) of subjects (e.g., records or record attributes) and the relationships between and among them.

When a record is exported from the live system, migrated in a system update, or transferred to the preserver, the attributes should be linked to the record and available to the user. When pulling together the data prior to export, the creator should also ensure that the data captured are the right data. For example, in the case of distribution lists, the creator must ensure that if the recipients specified on "List A" were changed at some point in the active life of records, the accurate "List A: Version 1" is exported with the records associated with the first version, and that the second version is sent forward with those records sent to recipients on "List A: Version 2."

A.2 Access Privileges

Defining access privileges means assigning responsibility for the creation, modification, annotation, relocation, and destruction of records on the basis of competence, which is the authority and capacity to carry out an administrative action. Implementing access privileges means conferring exclusive capability to exercise such responsibility. In electronic systems, access privileges are usually articulated in tables of user profiles. Effective implementation of access privileges involves the monitoring of access through an audit trail that records every interaction that an officer has with each record (with the possible exception of viewing the record). If the access privileges are not embedded within the electronic system but are based on an external security system (such as the exclusive assignment of keys to a location), the effective implementation of access privileges will involve monitoring the security system.

A.3 Protective Procedures: Loss and Corruption of Records

Procedures to protect records against loss or corruption include: prescribing regular back-up copies of records and their attributes; maintaining a system back-up that includes system programs, operating system files, etc.; maintaining an audit trail of additions and changes to records since the last periodic back-up; ensuring that, following any system failure, the back-up and recovery procedures will automatically guarantee that all complete updates (records and any control information such as indexes required to access the records) contained in the audit trail are reflected in the rebuilt files and also guarantee that any incomplete operation is backed up. The capability should be provided to rebuild forward from any back-up copy, using the back-up copy and all subsequent audit trails.

A.4 Protective Procedures: Media and Technology

Procedures to counteract media fragility and technological obsolescence include: planning upgrades to the organization's technology base; ensuring the ability to retrieve, access, and use

[16] If the attribute values contained in the profile are also expressed independently as entries in a register of all records made or received by the creator, then, in addition to establishing the identity and supporting the inference of the integrity of the record, they would corroborate such identity and strengthen the inference of integrity.

stored records when components of the electronic system are changed; refreshing the records by regularly moving them from one storage medium to another; and migrating records from an obsolescent technology to a new technology.

A.5 Establishment of Documentary Forms

The documentary form of a record may be determined in connection to a specific administrative procedure, or in connection to a specific phase(s) within a procedure. The documentary form may be prescribed by business process and work-flow control technology, where each step in an administrative procedure is identified by specific record forms. If a creator customizes a specific application, such as an electronic mail application, to carry certain fields, the customized form becomes, by default, the required documentary form. It is understood that the creator, acting either on the basis of its own needs or the requirements of the juridical system, not an individual officer, establishes the required documentary form(s) of records.

When the creator establishes the documentary form in connection to a procedure, or to specific phases of a procedure, it is understood that this includes the determination of the intrinsic and extrinsic elements of form[17] that will allow for the maintenance of the authenticity of the record. Because, generally speaking, that determination will vary from one form of a record to another, and from one creator to another, it is not possible to predetermine or generalize the relevance of specific intrinsic and extrinsic elements of documentary form in relation to authenticity.

A.6 Authentication of Records

In common usage, to authenticate means to prove or serve to prove the authenticity of something. More specifically, the term implies establishing genuineness by adducing legal or official documents or expert opinion. For the purposes of the benchmark requirements, authentication is understood to be a declaration of a record's authenticity at a specific point in time by a juridical person entrusted with the authority to make such declaration. It takes the form of an authoritative statement (which may be in the form of words or symbols) that is added to or inserted in the record attesting that the record is authentic.[18] The requirement may be met by linking the authentication of specific types of records to business procedures and assigning responsibility to a specific office or officer for authentication.

The authentication of copies differs from the validation of the process of reproduction of the digital components of the records. The latter process occurs every time the records of the creator are moved from one medium to another or migrated from one technology to another.

A.7 Identification of Authoritative Record

An authoritative record is a record that is considered by the creator to be its official record and is usually subject to procedural controls that are not required for other copies. The identification of authoritative records corresponds to the designation of an office of primary responsibility as one of the components of a record retention schedule. The Office of Primary Responsibility is the office given the formal competence for maintaining the authoritative (that

---

[17] The extrinsic and intrinsic elements of form are defined and explained in the InterPARES 1 *Template for Analysis* (see Authenticity Task Force (2000), "Appendix 1: Template for Analysis," in Duranti, *Long-term Preservation*, op. cit., 192–203. Online reprint available at http://www.interpares.org/book/interpares_book_j_app01.pdf).

[18] The meaning of authentication as it is used by the Authenticity Task Force in this report is broader than its meaning in public key infrastructure (PKI) applications. In such applications, authentication is restricted to proving identity and public key ownership over a communication network.

is, official) records belonging to a given class within an integrated classification scheme and retention schedule. The purpose of designating an office of primary responsibility for each class of record is to reduce duplication and to designate accountability for records.

It is understood that in certain circumstances there may be multiple authoritative copies of records, depending on the purpose for which the record is created.

A.8 Removal and Transfer of Relevant Documentation

This requirement implies that the creator needs to carry forward with the removed records all the information that is necessary to establish the identity and demonstrate the integrity of those records, as well as the information necessary to place the records in their relevant contexts.

APPENDIX 21B

Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records

Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records[1]

Preamble

The baseline requirements outline the minimum conditions necessary to enable the preserver to attest to the authenticity of copies of inactive electronic records.

Baseline Requirements (Requirement Set B)

The preserver should be able to demonstrate that:

REQUIREMENT B.1: Controls over Records Transfer, Maintenance, and Reproduction	the procedures and system(s) used to transfer records to the archival institution or program; maintain them; and reproduce them embody adequate and effective controls to guarantee the records' identity and integrity, and specifically that **B.1.a** Unbroken custody of the records is maintained; **B.1.b** Security and control procedures are implemented and monitored; and **B.1.c** The content of the record and any required annotations and elements of documentary form remain unchanged after reproduction.

REQUIREMENT B.2: Documentation of Reproduction Process and its Effects	the activity of reproduction has been documented, and this documentation includes **B.2.a** The date of the records' reproduction and the name of the responsible person; **B.2.b** The relationship between the records acquired from the creator and the copies produced by the preserver; **B.2.c** The impact of the reproduction process on their form, content, accessibility and use; and **B.2.d** In those cases where a copy of a record is known not to fully and faithfully reproduce the elements expressing its identity and integrity, such information has been documented by the preserver, and this documentation is readily accessible to the user;

[1] Excerpted from: Authenticity Task Force (2002), "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 204-219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf.

REQUIREMENT B.3: Archival Description	the archival description of the fonds containing the electronic records includes—in addition to information about the records' juridical-administrative, provenancial, procedural, and documentary contexts—information about changes the electronic records of the creator have undergone since they were first created.

Commentary on the Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records

The establishment and implementation of the baseline requirements take place as part of the function of managing preservation. The preservation function and the role of the baseline requirements within it are described in more detail in the "Preservation Task Force Report."

B.1 Controls over Records Transfer, Maintenance, and Reproduction

The controls over the transfer of electronic records to archival custody include establishing, implementing, and monitoring procedures for registering the records' transfer; verifying the authority for transfer; examining the records to determine whether they correspond to the records that are designated in the terms and conditions governing their transfer; and accessioning the records.

As part of the transfer process, the assessment of the authenticity of the creator's records, which has taken place as part of the appraisal process, should be verified. This includes verifying that the attributes relating to the records' identity and integrity have been carried forward with them (Requirement A.1), along with any relevant documentation (Requirement A.8).

The controls over the maintenance of electronic records once they have been transferred to archival custody are similar to several of the ones enumerated in the benchmark requirements. For example, the preserver should establish access privileges concerning the access, use, and reproduction of records (Requirement A.2); establish procedures to prevent, discover, and correct loss or corruption of records (Requirement A.3), as well as procedures to guarantee the continuing identity and integrity of records against media deterioration and across technological change (Requirement A.4). Once established, the privileges and procedures should be effectively implemented and regularly monitored. If authentication of the records is required, the preserver should establish specific rules regarding who is authorized to authenticate them and the means of authentication that will be used (Requirement A.6).

The controls over the reproduction of records include establishing, implementing, and monitoring reproduction procedures that are capable of ensuring that the content of the record is not changed in the course of reproduction.

B.2 Documentation of Reproduction Process and its Effects

Documenting the reproduction process and its effects is an essential means of demonstrating that the reproduction process is transparent (i.e., free from pretence or deceit). Such transparency is necessary to the effective fulfilment of the preserver's role as a trusted custodian of the records. Documenting the reproduction process and its effects is also important for the users of records since the history of reproduction is an essential part of the history of the record itself. Documentation of the process and its effects provides users of the records with a critical tool for assessing and interpreting the records.

B.3 Archival Description

Traditionally it has been a function of archival description to authenticate the records and perpetuate their administrative and documentary relationships. With electronic records, this function becomes critical. Once the records no longer exist except as authentic copies, the archival description is the primary source of information about the history of the record, that is, its various reproductions and the changes to the record that have resulted from them. Although it is true that the documentation of each reproduction of the record copies[2] may be preserved, the archival description summarizes the history of all the reproductions, thereby obviating the need to preserve all the documentation for each and every reproduction. In this respect, the description constitutes a collective attestation of the authenticity of the records and their relationships in the context of the fonds to which the records belong. This is different from a certificate of authenticity, which attests to the authenticity of individual records. The importance of this collective attestation is that it authenticates and perpetuates the relationships between and among records within the same fonds.

[2] Although, technically, every reproduction of a record that follows its acquisition by the preserver is an authentic copy, it is the only record that exists and, therefore, should normally be referred to as "the record" rather than as "the copy."

Appendix 21c

Digital Records Maintenance and Preservation Strategies[1]

This appendix includes a list of preservation strategies largely drawn from the UNESCO *Guidelines for the Preservation of Digital Heritage*,[2] which offers a framework for describing digital records preservation strategies that protect and maintain the accessibility of authentic copies of digital records throughout the chain of preservation.

The complete list of possible strategies adopted by InterPARES 2 is conceptually divided into two broad categories: a) maintenance strategies and b) preservation strategies.

A. Maintenance Strategies

Maintenance strategies[3] are the minimum necessary requirement to protect and maintain accessibility of authentic copies of digital records. There are eight primary maintenance strategies. All are necessary to ensure the records components will exist long enough for preservation strategies to be applied.

A1. Clear allocation of responsibilities

A person or office must be given unambiguous responsibility for managing records storage and protection. This is a technical responsibility that requires a specific skill set, dedicated resources, and an appropriate plan. This strategy can be undertaken by hiring a competent staff member devoted exclusively to this task or by assigning existing staff or an existing office a specific portion of time to carry out the responsibilities.

A2. Provision of the appropriate technical infrastructure

This includes all of the physical and administrative resources that enable the recordkeeping and/or maintenance processes (buildings, computer hardware, computer networks and the auxiliary staff necessary to maintain the same).

A3. System maintenance, support and replacement

The implementation of a plan for maintaining, updating and/or replacing hardware and software.

A4. Transfer of data to new storage media on a regular basis

The implementation of a plan for copying of data from one storage medium to another to avoid the impact of media decay. Such transfers should be undertaken in a systematic manner.

A5. Adherence to appropriate conditions for storage media

The rate of media decay may be dramatically reduced by adhering to appropriate environmental conditions. For instance, excessive heat, humidity and dust endanger storage media.

---

[1] Adapted from: Kevin Glick, "Electronic Records Preservation Strategies," (unpublished report, 2006).

[2] Colin Webb (2003), *Guidelines for the Preservation of Digital Heritage*. Prepared by the National Library of Australia for the Information Society Division, United Nations Educational, Scientific and Cultural Organization, report no. CI-2003/WS/3. Available at http://unesdoc.unesco.org/images/0013/001300/130071e.pdf.

[3] A maintenance strategy is defined as "A coherent set of objectives and methods for protecting (i.e., safeguarding authenticity and ensuring accessibility of) digital components and related information over time while still in active or semi-active use by the creator, and for reproducing the related authentic records and/or record aggregations" (InterPARES 2 Terminology Database, op. cit.).

A6. Redundancy and geographic location

The duplication of digital objects and the storage of the resulting multiple copies on different physical media protects them against media failure. Storage in different physical locations protects against poor environmental storage conditions, fire, flood, etc., at a particular storage site.

A7. System security

Controls should be implemented to ensure that digital components of records are exposed only to authorized users and/or processes. Such controls should include restricting physical access to places where computers are kept as well as restricting access to the digital records on the computers themselves. The latter can be accomplished through various means, including the use of passwords and/or biometric authentication to log on to the system.

A8. Disaster planning

The strategies listed above are designed to minimize accidental loss of data and maximize media longevity, but even with perfect storage conditions and excellent handling protocols, disasters may still happen. A disaster recovery plan should contain detailed procedures for restoring a damaged system and for guiding the effective recovery of recordkeeping and/ or preservation systems following a disaster.

B. Preservation Strategies

In addition to the maintenance strategies, every records preserver is responsible for establishing a trusted preservation system[4] for expressing one or more preservation strategies.[5] Twelve preservation strategies are listed below, in Section B, divided into four broadly defined groups. It is most likely that, in practice, a preserver will support two or more preservation strategies in addition to the eight maintenance strategies listed above in Section A.

B1. Use of standards

The use of widely available and supported standards increases the likelihood of stability and longer term support. Such standards may either be *de jure*,[6] if they have been formally agreed upon, or *de facto*,[7] if they have been widely adopted by industry. Standards can apply to many facets of a preservation system, including encoding methods, file formats, physical storage media, etc. Compliance with standards might also simplify the application and/or maximize the effectiveness of later preservation strategies. Standardization may be applied *prospectively*, by limiting the formats in which digital records may be transferred to the preserver; or *retrospectively*, by converting files received in other formats to standard ones.

---

[4] A trusted preservation system is defined as "The whole of the rules that control the preservation and use of the records of the creator and provide a circumstantial probability of the authenticity of the records, and the tools and mechanisms used to implement those rules" (Ibid.).

[5] A preservation strategy is defined as "A coherent set of objectives and methods for protecting (i.e., safeguarding authenticity and ensuring accessibility of) digital components and related information of inactive records over time, and for reproducing the related authentic records and/or archival aggregations" (Ibid.).

[6] A *de jure* standard is defined as a "Standard issued by an official standards-setting body, whether national (e.g., ANSI), multi-national (e.g., CEN) or international (e.g., ISO)" (Ibid.). For computer file formats, two recent *de jure* standards are PDF/A (PDF standard for archiving) and ODF (OASIS OpenDocument Format).

[7] A *de facto* standard is defined as a "Standard not issued by any official standards-setting body, but nevertheless widely used and recognized by its users as a standard" (Ibid.). Well known and widely used computer file formats that are considered *de facto* standards include PDF, TIFF, DOC and ZIP.

B1.1. Self-describing formats (persistent object preservation, tagging)

Analysis and tagging of records so that the functions, relationships and structure of specific elements can be described. The re-presentation of content can be liberated from specific software applications and can be achieved using different applications as technology changes.

B1.2. Encapsulation

Binding together a record and the means of providing access to it, normally in a *wrapper* that describes what it is in a way that can be understood by a wide range of technologies (such as an XML document). The wrapper often includes metadata that describe or link to the correct tools.

B1.3. Restricting the range of formats to be managed (normalization)

Storing records in a limited number of formats only.[8] The selection of acceptable formats may continue to include new proprietary formats or new generations of existing proprietary formats, or it may be restricted to non-proprietary formats, to carry standardization one step further. One example of this approach is referred to as *durable encoding*, which recommends encoding records to conform to well-known data processing standards down to the level of encoding bits as ASCII or Unicode UTF-8, and objects as XML.

B1.4. Conversion

Transferring digital records from one hardware or software generation to another. As distinct from *refreshing*, which copies the data stream from one carrier to another, conversion entails transforming the logical form of a digital object so that the conceptual object can continue to be correctly rendered or presented by the new hardware or software. The most commonly proposed conversion method involves permanently transforming one logical format into another in line with technological change, so that all converted objects can be presented with prevailing technology. It is also possible to propose a “conversion on demand” or “conversion at the point of access” model. This approach is discussed below in Section B2.4.

B2. Technology dependence

These strategies continue to rely on the original hardware and/or software without changing the records.

B2.1. Technology preservation

Maintaining the original software and hardware with which digital records were presented.

B2.2. Reliance on backward compatibility

Trusting the ability of some software to correctly interpret and present digital components of records created with previous versions of the same software. In this strategy, the presentation is limited to a temporary conversion for viewing or for non-archival copying purposes, whereas conversion permanently changes records into the format supported by the current version of the software.

---

[8] For a detailed analysis of current issues and trends in the selection of file, wrapper, tagging and encoding formats, together with recommendations for developing and implementing policies on selecting digital file formats for long-term preservation, see: Evelyn Peters McLellan (2006), “InterPARES 2 Project - General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation.” Available in English at http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_english.pdf, and in French at http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_french.pdf.

B2.3. Software re-engineering

Transforming software as technologies change. As such, it is similar to the transformation of record formats, discussed in sections B1.4.and B2.2. This may include anything from re-compiling source code for a new platform to re-coding the software from scratch in another programming language.

B2.4. Viewers and conversion at the point of access

The use of software tools or transformation methods that provide temporary accessibility when needed, using the original data stream.

B2.5. Emulation

Using software that makes one technology behave like another. In other words, making future technologies behave like the original environment of a preserved digital record, so that the original record could be presented in its original manifestation from the original, or converted, data streams.

B3. Non-digital approaches

Copying the digital records onto relatively stable analogue media, such as paper or microfilm; shifting the preservation burden to an analogue copy in place of the digital object. This approach destroys any functionality provided by the software, such as manipulability.

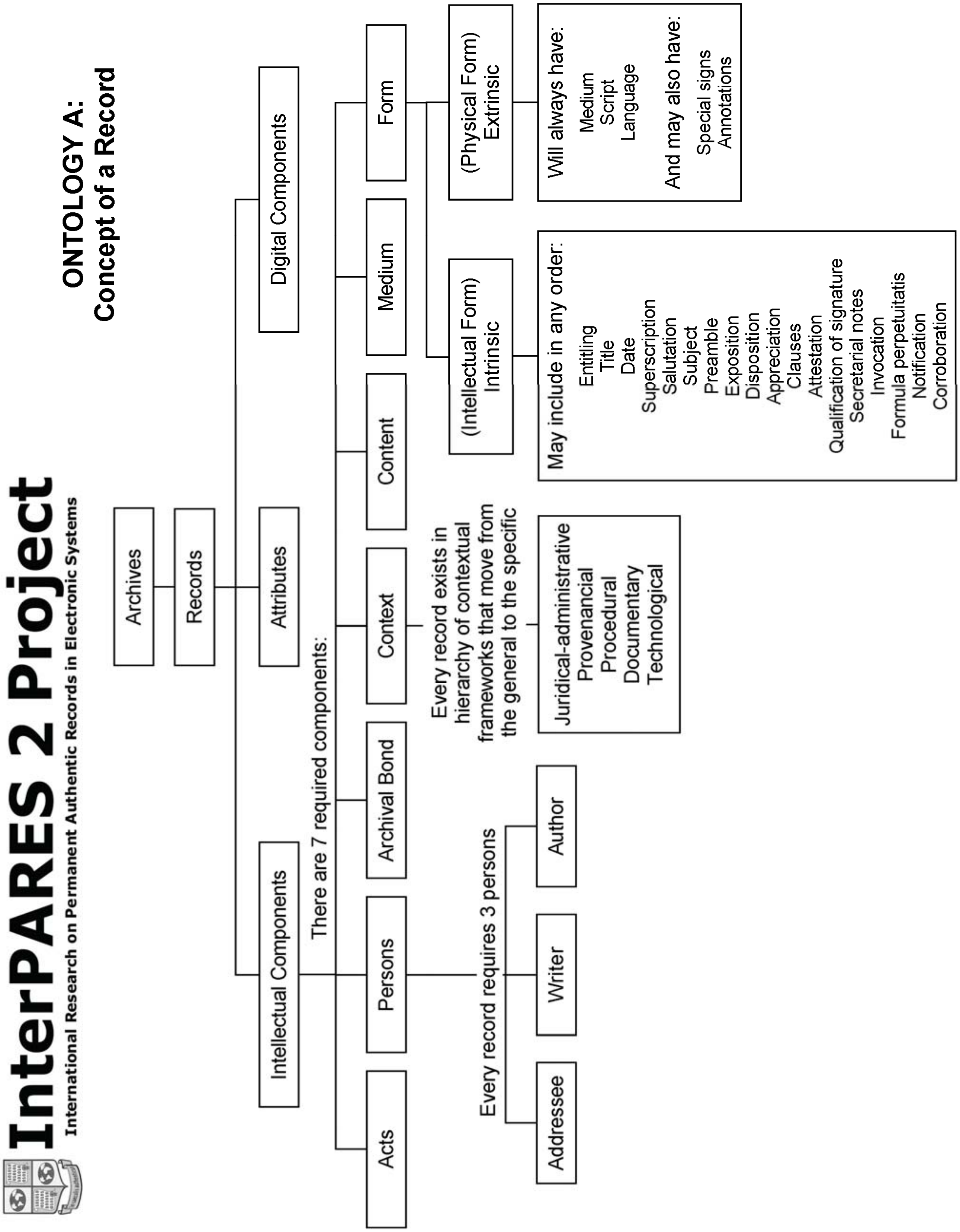
B4. Data restoration (digital archaeology)

Recovering records as bits from physical media followed by steps to restore the intelligibility of the recovered records. It is most often employed in the recovery of data from failed, damaged or degraded media, but methods to restore intelligibility have been used to rescue documents in obsolete formats.

APPENDIX 22

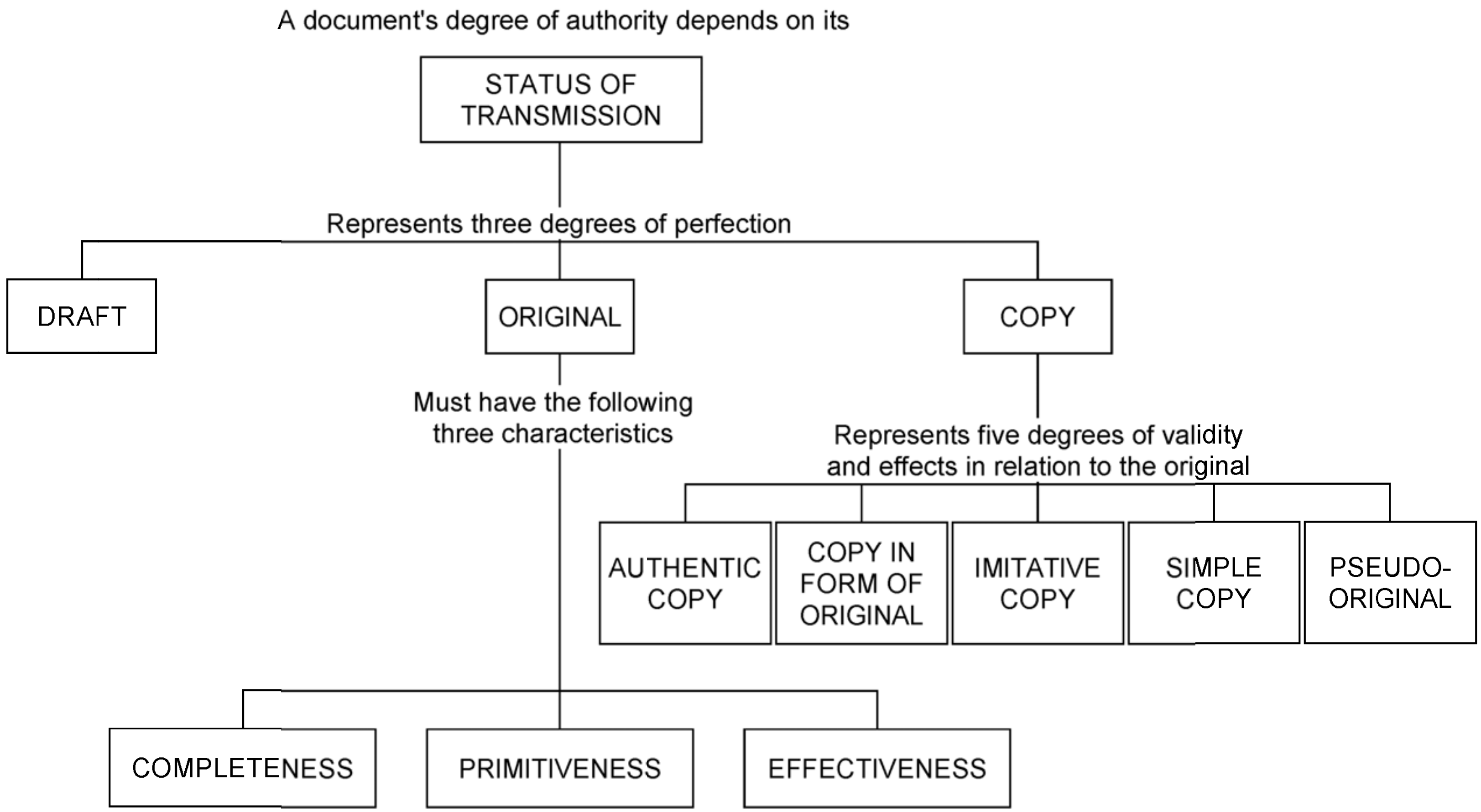
InterPARES 2 Project Ontologies

Ontology A: Concept of a Record
Ontology B: Concept of the Status of Transmission of a Record
Ontology C: Trustworthiness of a Record


InterPARES 2 Project
International Research on Permanent Authentic Records in Electronic Systems
ONTOLOGY A:
Concept of a Record
Archives
Records
Attributes
Intellectual Components
Digital Components
There are 7 required components:
Acts
Persons
Archival Bond
Context
Content
Medium
Form
Every record requires 3 persons
Addressee
Writer
Author
Every record exists in hierarchy of contextual frameworks that move from the general to the specific
Juridical-administrative
Provenancial
Procedural
Documentary
Technological
(Intellectual Form)
Intrinsic
(Physical Form)
Extrinsic
May include in any order:
Entitling
Title
Date
Superscription
Salutation
Subject
Preamble
Exposition
Disposition
Appreciation
Clauses
Attestation
Qualification of signature
Secretarial notes
Invocation
Formula perpetuitatis
Notification
Corroboration
Will always have:
Medium
Script
Language
And may also have:
Special signs
Annotations

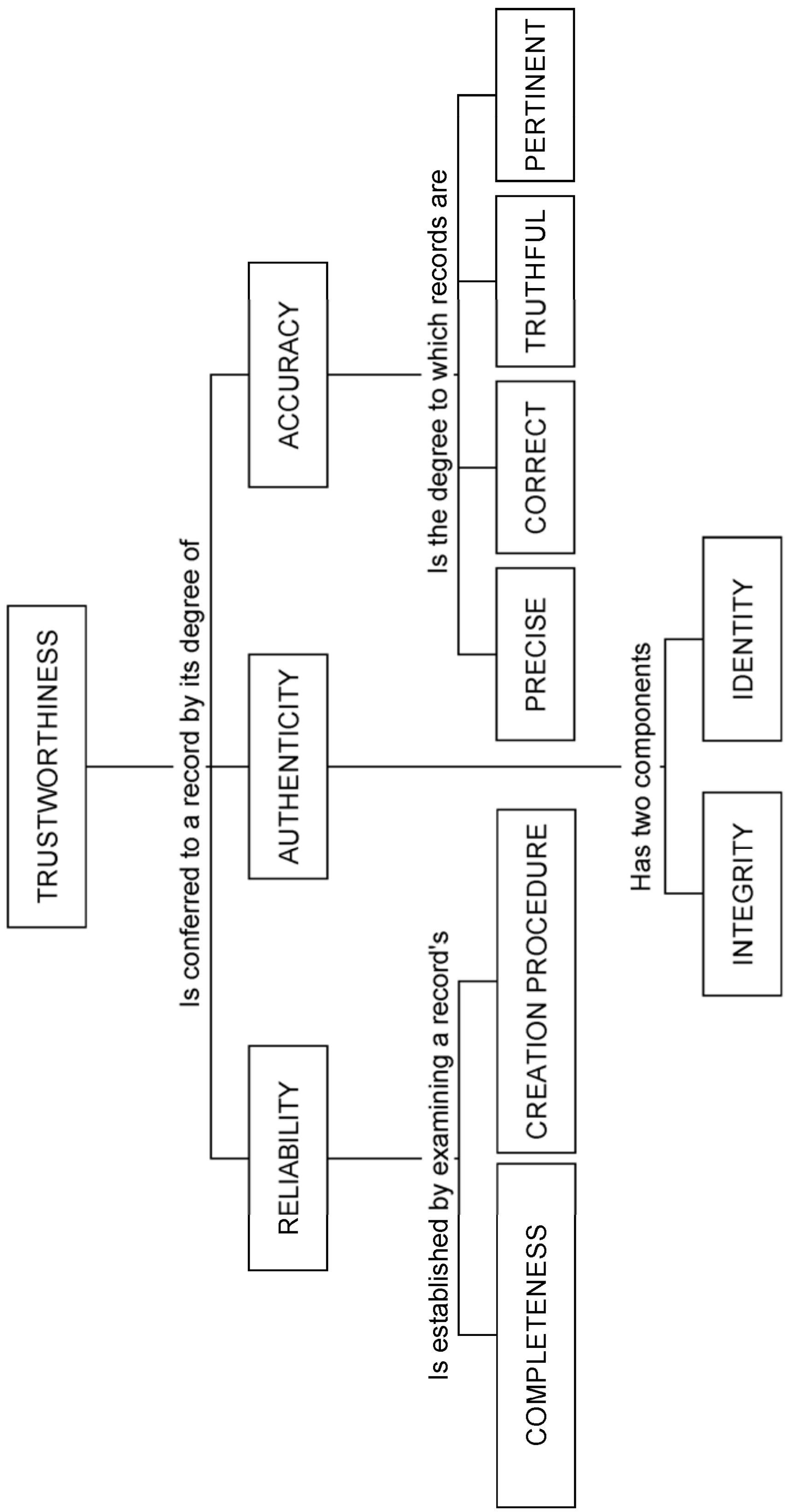
InterPARES 2 Project
International Research on Permanent Authentic Records in Electronic Systems

ONTOLOGY B:
Concept of the Status of Transmission of a Record



InterPARES 2 Project
International Research on Permanent Authentic Records in Electronic Systems

ONTOLOGY C:
Trustworthiness of a Record



APPENDIX 23

Select Bibliography of InterPARES 2 Project Documents

Select Bibliography of InterPARES 2 Project Documents

This appendix lists most of the key InterPARES 2 Project documents, arranged by Project unit. Published documents and documents related to specific Project dissemination activities are not included.[1] Copies of all listed documents and reports are available on the InterPARES 2 Web site and on the DVD that accompanies this book.

Focuses 1, 2 and 3

Hackett, Yvette, William Underwood and Philip Eppard (2008), "InterPARES 2 Project - Focus Task Force Report: Case and General Studies in the Artistic, Scientific and Governmental Sectors." Available in Part 1 and in electronic form at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_1_focus_task_force.pdf.

Focus 1: Artistic Activities

Case Studies

The following is a list of the key documents, beginning with the final report, for each case study in Focus 1 arranged numerically by case study number.

01 Martine Cardin (2004), "InterPARES 2 Project - Case Study 01 Final Report: Arbo Cyber, théâtre (?)."
http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_study_english.pdf
http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_study_french.pdf

- Case Study Proposal (Martine Cardin, 2002)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs01_proposal.pdf
- Characterization (Heather Daly, 2004)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs01_characterization.pdf
- Areas that Should Be Covered (Heather Daly, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs01_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs01_overview.pdf
- Diplomatic Analysis (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs01_diplomatic_analysis.pdf
- Domain 1 Research Questions (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs01_d1_questions.pdf

[1] See the Project's online Dissemination Database (http://www.interpares.org/ip2/ip2_dissemination.cfm?proj=ip2) for information about InterPARES 2 Project dissemination-related documents, including conference papers, book chapters, lectures, workshops, seminars, etc.

- Domain 3 Research Analysis (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs01_d3_analysis.pdf

02 Henry Daniel and Cara Payne (2004), "InterPARES 2 Project - Case Study 02 Final Report: Performance Artist Stelarc."
http://www.interpares.org/display_file.cfm?doc=ip2_cs02_final_report.pdf

- Case Study Proposal (Henry Daniel, 2002)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs02_proposal.pdf
- Characterization (Heather Daly, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs02_characterization.pdf
- Areas that Should Be Covered (Heather Daly and Ann Forman, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs02_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs02_overview.pdf
- Diplomatic Analysis (Tracey Krause, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs02_diplomatic_analysis.pdf
- Domain 1 Research Questions (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs02_d1_questions.pdf
- Domain 3 Research Analysis (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs02_d3_analysis.pdf
- Activity Model (see Case Study 02 Final Report)
- Entity Relationship Model (see Case Study 02 Final Report)

03 Brent Lee (2004), "InterPARES 2 Project - Case Study 03 Final Report: *HorizonZero*/ZeroHorizon Online Magazine and Media Database."
http://www.interpares.org/display_file.cfm?doc=ip2_cs03_final_report.pdf

- Case Study Proposal (Susan Kennard and Brent Lee, 2002)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs03_proposal.pdf
- Characterization (Ann Forman, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs03_characterization.pdf
- Areas that Should Be Covered (Ann Forman, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs03_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs03_overview.pdf
- Diplomatic Analysis (Tracey Krause, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs03_diplomatic_analysis.pdf
- Domain 1 Research Questions (Natalie Catto, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs03_d1_questions.pdf
- Domain 3 Research Analysis (Tracey Krause, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs03_d3_analysis.pdf

09(01) Isabella Orefice (2004), "InterPARES 2 Project - Case Study 09(1) Final Report: Digital Moving Images - Altair4 di Roma. A Multimedia Archaeological Project: *The House of Julius Polybius*."
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-1_final_report.pdf

- Case Study Proposal (Marta Braun et al., 2001)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09_proposal.pdf
- Characterization (Heather Daly, 2005)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-1_characterization.pdf
- Areas that Should Be Covered (Heather Daly, 2005)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-1_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-1_overview.pdf
- Diplomatic Analysis (Natalie Catto, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-1_diplomatic_analysis.pdf
- Domain 1 Research Questions (Natalie Catto, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-1_d1_questions.pdf
- Domain 3 Research Analysis (Natalie Catto, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-1_d3_analysis.pdf
- Activity Model (see Case Study 09-1 Final Report)

09(02) Andrew Rodger (2006), "InterPARES 2 Project - Case Study 09(2) Final Report: Digital Moving Images - National Film Board of Canada."
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-2_final_report.pdf

- Case Study Proposal (Marta Braun et al., 2001)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09_proposal.pdf
- Characterization (Geneviève Sheppard, 2005)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-2_characterization.pdf
- Areas that Should Be Covered (Geneviève Sheppard, 2005)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-2_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-2_overview.pdf
- Diplomatic Analysis (Natalie Catto, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-2_diplomatic_analysis.pdf
- Domain 1 Research Questions (Geneviève Sheppard, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-2_d1_questions.pdf
- Domain 3 Research Analysis (Natalie Catto, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-2_d3_analysis.pdf

09(03) James M. Turner et al. (2004), "InterPARES 2 Project - Case Study 09(3) Final Report: Digital Moving Images - Commercial Film Studio."
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_final_report.pdf

- Case Study Proposal (Marta Braun et al., 2001)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs09_proposal.pdf
- Characterization (Geneviève Sheppard, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_characterization.pdf
- Areas that Should Be Covered (Geneviève Sheppard, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_overview.pdf
- Diplomatic Analysis (Tracey Krause, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_diplomatic_analysis.pdf
- Domain 1 Research Questions (Geneviève Sheppard, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_d1_questions.pdf
- Domain 3 Research Analysis (Tracey Krause, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_d3_analysis.pdf

09(04) Mary Ide (2005), InterPARES 2 Project - Case Study 09(4) Final Report: Digital Moving Images - WGBH Boston."
http://www.interpares.org/display_file.cfm?doc=ip2_cs09-4_final_report.zip

- Case Study Proposal (Marta Braun et al., 2001)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs09_proposal.pdf
- Characterization (Geneviève Sheppard, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs09-4_characterization.pdf
- Areas that Should Be Covered (Geneviève Sheppard, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs09-4_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs09-4_overview.pdf
- Diplomatic Analysis (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs09-4_diplomatic_analysis.pdf
- Domain 1 Research Questions (Geneviève Sheppard, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs9-4_d1_questions.pdf
- Domain 3 Research Analysis (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs9-4_d3_analysis.pdf

10 Sally Hubbard (2006), "InterPARES 2 Project - Case Study 10 Final Report: *The Danube Exodus*." http://www.interpares.org/display_file.cfm?doc=ip2_cs10_final_report.pdf

- Case Study Proposal (Sally Hubbard, 2002) http://www.interpares.org/display_file.cfm?doc=ip2_cs10_proposal.pdf
- Characterization (Geneviève Sheppard, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_cs10_characterization.pdf
- Areas that Should Be Covered (Geneviève Sheppard, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_cs10_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_cs10_overview.pdf
- Diplomatic Analysis (Cobi Falconer, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_cs10_diplomatic_analysis.pdf
- Domain 1 Research Questions (Jennifer Douglas, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_cs10_d1_questions.pdf
- Domain 3 Research Analysis (Cobi Falconer, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_cs10_d3_analysis.pdf

13 J. Scott Amort (2004), "InterPARES 2 Project - Case Study 13 Final Report: *Obsessed Again….*" http://www.interpares.org/display_file.cfm?doc=ip2_cs13_final_report.pdf

- Case Study Proposal (Keith Hamel, 2002) http://www.interpares.org/display_file.cfm?doc=ip2_cs13_proposal.pdf
- Characterization (Ann Forman, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_cs13_characterization.pdf
- Areas that Should Be Covered (Ann Forman, 2005) http://www.interpares.org/display_file.cfm?doc=ip2_cs13_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_cs13_overview.pdf
- Diplomatic Analysis (Melissa Adams and Tracey Krause, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_cs13_diplomatic_analysis.pdf
- Domain 1 Research Questions (Jennifer Douglas, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_cs13_d1_questions.pdf
- Domain 3 Research Analysis (Tracey Krause, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_cs13_d3_analysis.pdf
- Activity Model (see Case Study 13 Final Report)
- Entity Relationship Model (see Case Study 13 Final Report)

15 Sydney Fels and Seth Dalby (2004), “InterPARES 2 Project - Case Study 15 Final Report: *Waking Dream*.”
http://www.interpares.org/display_file.cfm?doc=ip2_cs15_final_report.pdf

- Case Study Proposal (Sydney Fels, 2002)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs15_proposal.pdf
- Characterization (Heather Daly and Ann Forman, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs15_characterization.pdf
- Areas that Should Be Covered (Heather Daly, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs15_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs15_overview.pdf
- Diplomatic Analysis (Natalie Catto, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs15_diplomatic_analysis.pdf
- Domain 1 Research Questions (Jennifer Douglas, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs15_d1_questions.pdf
- Domain 3 Research Analysis (Tracey Krause, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs15_d3_analysis.pdf
- Activity Model (see Case Study 15 Final Report)
- Entity Relationship Model (see Case Study 15 Final Report)

22 Hafner, Nadine, Janine Johnston, Tracey Krause and Keum Hee Yu (2006), “InterPARES 2 Project - Case Study 22 Final Report: Electronic Café International (ECI).” (retired)
http://www.interpares.org/display_file.cfm?doc=ip2_cs22_final_report_DRAFT.pdf

- Case Study Proposal (Howard Bessar, 2003)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs22_proposal.pdf
- Characterization (Geneviève Sheppard, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs21_characterization_UNVALIDATED.pdf
- Areas that Should Be Covered (Geneviève Sheppard, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs22_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs22_overview.pdf
- Diplomatic Analysis (Nadine Hafner, Keum Hee and Tracey Krause, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs22_diplomatic_analysis.pdf
- Domain 1 Research Questions (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs22_d1_questions.pdf
- Domain 3 Research Analysis (Tracey Krause, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs22_d3_analysis.pdf

GENERAL STUDIES

03 Jennifer Douglas (2006), "InterPARES 2 Project - General Study 03 Final Report: Preserving Interactive Digital Music - The MUSTICA Initiative." http://www.interpares.org/display_file.cfm?doc=ip2_gs03_final_report.pdf

John Roeder (2006), "InterPARES 2 Project - General Study 03: Authenticity of digital music: key insights from interviews in the MUSTICA project." http://www.interpares.org/display_file.cfm?doc=ip2_gs03_authenticity_roeder_v2.pdf

- Overview (Peter Gagné, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_gs03_overview.pdf

04 Michael Longton (2004), "InterPARES 2 Project - General Study 04 Final Report: Recordkeeping Practices of Composers." http://www.interpares.org/display_file.cfm?doc=ip2_gs04_final_report.pdf

- Overview (Peter Gagné, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_gs04_overview.pdf

05 Michèle V. Cloonan and Shelby Sanett (2004), "InterPARES 2 Project - General Study 05 Final Report, Round 2: The Preservation of Authentic Electronic Records: *Ad hoc*, Inconsistent, or Strategic?" http://www.interpares.org/display_file.cfm?doc=ip2_gs05_r2_final_report.pdf

07 Jessica Bushey and Marta Braun (2006), "InterPARES 2 Project - GS07 Final Report: Survey of Recordkeeping Practices of Photographers using Digital Technology." http://www.interpares.org/display_file.cfm?doc=ip2_gs07_final_report.pdf

- Overview (Peter Gagné, 2006) http://www.interpares.org/display_file.cfm?doc=ip2_gs07_overview.pdf

Focus 2: Scientific Activities

CASE STUDIES

The following is a list of the key documents, beginning with the final report, for each case study in Focus 2 arranged numerically by case study number.

06 Tracey P. Lauriault and Yvette Hackett (2005), "InterPARES 2 Project - Case Study 06 Final Report: Cybercartographic Atlas of Antarctica." http://www.interpares.org/display_file.cfm?doc=ip2_cs06_final_report.zip

- Case Study Proposal (Tracey P. Lauriault, 2002) http://www.interpares.org/display_file.cfm?doc=ip2_cs06_proposal.pdf
- Characterization (Geneviève Sheppard, 2005) http://www.interpares.org/display_file.cfm?doc=ip2_cs06_characterization.pdf
- Areas that Should Be Covered (Geneviève Sheppard, 2005) http://www.interpares.org/display_file.cfm?doc=ip2_cs06_areas_2b_covered.pdf

- Overview (Peter Gagné, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs06_overview.pdf
- Diplomatic Analysis (Sherry Xie, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs06_diplomatic_analysis.pdf
- Domain 1 Research Questions (Sherry Xie, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs06_d1_questions.pdf
- Domain 3 Research Analysis (Sherry Xie, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs06_d3_analysis.pdf

08 William Underwood (2005), "InterPARES 2 Project - Case Study 08 Final Report: Mars Global Surveyor Data Records in the Planetary Data System."
http://www.interpares.org/display_file.cfm?doc=ip2_cs08_final_report.pdf

- Case Study Proposal (William Underwood, 2002)
http://www.interpares.org/display_file.cfm?doc=ip2_cs08_proposal.pdf
- Characterization (Geneviève Sheppard, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs08_characterization.pdf
- Areas that Should Be Covered (Geneviève Sheppard, 2005)
http://www.interpares.org/display_file.cfm?doc=ip2_cs08_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs08_overview.pdf
- Diplomatic Analysis (unavailable)
- Domain 1 Research Questions (Carolyn Petrie, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs08_d1_questions.pdf
- Domain 3 Research Analysis (Carolyn Petrie, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs08_d3_analysis.pdf

14 Richard Pearce-Moses, Erin O'Meara and Randy Preston (2004), "InterPARES 2 Project - Case Study 14 Final Report: Archaeological Records in a Geographical Information System: Research in the American Southwest."
http://www.interpares.org/display_file.cfm?doc=ip2_c14_final_report.pdf

- Case Study Proposal (Erin O'Meara, 2003)
http://www.interpares.org/display_file.cfm?doc=ip2_cs14_proposal.pdf
- Characterization (Ann Forman, 2005)
http://www.interpares.org/display_file.cfm?doc=ip2_cs14_characterization.pdf
- Areas that Should Be Covered (Heather Daly and Ann Forman, 2005)
http://www.interpares.org/display_file.cfm?doc=ip2_cs14_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs14_overview.pdf
- Diplomatic Analysis (Natalie Catto, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs14_diplomatic_analysis.pdf

- Domain 1 Research Questions (Natalie Catto, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs14_d1_questions.pdf
- Domain 3 Research Analysis (Natalie Catto, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs14_d3_analysis.pdf
- Activity Model (see Case Study 14 Final Report)

19 Kenneth Hawkins (2006), "InterPARES 2 Project - Case Study 19 Final Report: Preservation and Authentication of Electronic Engineering and Manufacturing Records."
http://www.interpares.org/display_file.cfm?doc=ip2_cs19_final_report.pdf

- Case Study Proposal (Kenneth Thibodeau, 2003)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs19_proposal.pdf
- Characterization (Geneviève Sheppard, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs19_characterization.pdf
- Areas that Should Be Covered (Geneviève Sheppard, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs19_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs19_overview.pdf
- Diplomatic Analysis (Kenneth Hawkins, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs19_diplomatic_analysis.pdf
- Domain 1 Research Questions (Geneviève Sheppard, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs19_d1_questions.pdf
- Domain 3 Research Analysis (Kenneth Hawkins and Yvette Hackett, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs19_d3_analysis.pdf
- Activity Model (see Case Study 19 Final Report)

26 Bart Ballaux (2005), "InterPARES 2 Project - Case Study 26 Final Report: Microvariability & Oscillations of Stars (MOST) Satellite Mission - Preservation of Space Telescope Data."
http://www.interpares.org/display_file.cfm?doc=ip2_cs26_final_report.pdf

- Case Study Proposal (Luciana Duranti, 2004)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs26_proposal.pdf
- Characterization (Seth Dalby, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs26_characterization.pdf
- Areas that Should Be Covered (Seth Dalby, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs26_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs26_overview.pdf
- Diplomatic Analysis (Sherry Xie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs26_diplomatic_analysis.pdf

- Domain 1 Research Questions (Natalie Catto, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs26_d1_questions.pdf
- Domain 3 Research Analysis (Sherry Xie, 2006)
http://www.interpares.org/display_file.cfm?doc=ip2_cs26_d3_analysis.pdf
- Activity Model (see Case Study 26 Final Report)

GENERAL STUDIES

01 Reagan W. Moore (2004), "InterPARES 2 Project - General Study 01 Final Report: Building Preservation Environments with Data Grid Technology."
http://www.interpares.org/display_file.cfm?doc=ip2_gs01_final_report.pdf

06 William Underwood and Sheila Isbell (2007), "InterPARES 2 Project - General Study 06 Final Report: A Bayesian Belief Network: Supporting the Assessment of the Degree of Belief that a Recordkeeping System Maintains Authentic Digital Records."
http://www.interpares.org/display_file.cfm?doc=ip2_gs06_final_report.pdf

09 Randy Preston (2006), "InterPARES 2 Project - General Study 09 Final Report: Digital Recordkeeping Practices of GIS Archaeologists Worldwide: Results of a Web-based Survey."
http://www.interpares.org/display_file.cfm?doc=ip2_gs09_final_report.pdf

10 Tracey P. Lauriault and Barbara L. Craig (2007), "InterPARES 2 Project - General Study 10 Final Report: Preservation Practices of Scientific Data Portals."
http://www.interpares.org/display_file.cfm?doc=ip2_gs10_final_report.zip

12 William Underwood, Kevin Glick and Mark Wolfe (2007), "InterPARES 2 Project - General Study 12 Final Report: Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data."
http://www.interpares.org/display_file.cfm?doc=ip2_gs12_final_report.pdf

Focus 3: Governmental Activities

CASE STUDIES

The following is a list of the key documents, beginning with the final report, for each case study in Focus 3 arranged numerically by case study number.

05 Jim Suderman et al. (2004), "InterPARES 2 Project - Case Study 05 Final Report: Archives of Ontario Web Exhibits."
http://www.interpares.org/display_file.cfm?doc=ip2_cs05_final_report.pdf

- Case Study Proposal (Jim Suderman, 2002)
http://www.interpares.org/display_file.cfm?doc=ip2_cs05_proposal.pdf
- Characterization (Heather Daly, 2005)
http://www.interpares.org/display_file.cfm?doc=ip2_cs05_characterization.pdf

- Areas that Should Be Covered (Heather Daly, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs05_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs05_overview.pdf
- Diplomatic Analysis (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs05_diplomatic_analysis.pdf
- Domain 1 Research Questions (Natalie Catto, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs05_d1_questions.pdf
- Domain 3 Research Analysis (Natalie Catto, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs05_d3_analysis.pdf
- Activity Model (see Case Study 05 Final Report)

12 Paul Arthur Berkman et al. (2005), "InterPARES 2 Project - Case Study 12 Final Report: Antarctic Treaty Searchable Database Case Study."
http://www.interpares.org/display_file.cfm?doc=ip2_cs12_final_report.pdf

- Case Study Proposal (Paul Arthur Berkman, 2002)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs12_proposal.pdf
- Characterization (Heather Daly, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs12_characterization.pdf
- Areas that Should Be Covered (Heather Daly, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs12_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs12_overview.pdf
- Diplomatic Analysis (Heather Daly, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs12_diplomatic_analysis.pdf
- Domain 1 Research Questions (Jennifer Douglas, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs12_d1_questions.pdf
- Domain 3 Research Analysis (Adam Farrell, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs12_d3_analysis.pdf
- Activity Model (see Case Study 12 Final Report)

17 Phil Eppard, Terry Maxwell and Mark Wolfe (2006), "InterPARES 2 Project - Case Study 17 Final Report: New York State Department of Motor Vehicles On-line Services System."
http://www.interpares.org/display_file.cfm?doc=ip2_cs17_final_report.pdf

- Case Study Proposal (Mark Wolfe, 2003)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs17_ proposal.pdf
- Characterization (Natalie Catto, 2007)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs17_characterization.pdf
- Areas that Should Be Covered (unavailable)

- Overview (unavailable)
- Diplomatic Analysis (Sherry Xie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs17_diplomatic_analysis.pdf
- Domain 1 Research Questions (unavailable)
- Domain 3 Research Analysis (unavailable)

18 Jean-François Blanchette, François Banat-Berger and Geneviève Shepherd (2004), "InterPARES 2 Project - Case Study 18 Final Report: Computerization of Alsace-Moselle's Land Registry."
http://www.interpares.org/display_file.cfm?doc=ip2_cs18_final_report.pdf

- Case Study Proposal (Jean-François Blanchette, 2003)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs18_proposal.pdf
- Characterization (Heather Daly, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs18_characterization.pdf
- Areas that Should Be Covered (Heather Daly, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs18_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs18_overview.pdf
- Diplomatic Analysis (Jennifer Douglas, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs18_diplomatic_analysis.pdf
- Domain 1 Research Questions (Geneviève Sheppard, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs18_d1_questions.pdf
- Domain 3 Research Analysis (Adam Farrell, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs18_d3_analysis.pdf
- Activity Model (see Case Study 18 Final Report)

20 John McDonough, Ken Hannigan and Tom Quinlan (2005), "InterPARES 2 Project - Case Study 20 Final Report: Revenue On-Line Service (ROS)."
http://www.interpares.org/display_file.cfm?doc=ip2_cs20_final_report.pdf

- Case Study Proposal (John McDonough, 2003)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs20_proposal.pdf
- Characterization (Seth Dalby, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs20_characterization.pdf
- Areas that Should Be Covered (Geneviève Sheppard, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs20_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs20_overview.pdf
- Diplomatic Analysis (Tracey Krause, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs20_diplomatic_analysis.pdf

- Domain 1 Research Questions (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs20_d1_questions.pdf
- Domain 3 Research Analysis (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs20_d3_analysis.pdf
- Activity Model (see Case Study 20 Final Report)

21 Elaine Goh (2005), "InterPARES 2 Project - Case Study 21 Final Report: The Electronic Filing System (EFS) of the Supreme Court of Singapore."
http://www.interpares.org/rws/display_file.cfm?doc=ip2_cs21_final_report.pdf

- Case Study Proposal (Elaine Goh, 2003)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs21_proposal.pdf
- Characterization (Geneviève Sheppard and Natalie Catto, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs21_characterization.pdf
- Areas that Should Be Covered (Geneviève Sheppard, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs21_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs21_overview.pdf
- Diplomatic Analysis (Natalie Catto, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs21_diplomatic_analysis.pdf
- Domain 1 Research Questions (Geneviève Sheppard and Natalie Catto, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs21_d1_questions.pdf
- Domain 3 Research Analysis (Natalie Catto, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs21_d3_analysis.pdf
- Activity Model (see Case Study 21 Final Report)

24 Evelyn McLellan (2005), "InterPARES 2 Project - Case Study 24 Final Report: City of Vancouver Geographic Information System (VanMap)."
http://www.interpares.org/display_file.cfm?doc=ip2_cs24_final_report.pdf

- Case Study Proposal (Evelyn Peters McLellan, 2004)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs24_proposal.pdf
- Characterization (Heather Daly, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs24_characterization.pdf
- Areas that Should Be Covered (Heather Daly, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs24_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs24_overview.pdf
- Diplomatic Analysis (Natalie Catto, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs24_diplomatic_analysis.pdf
- Domain 1 Research Questions (Natalie Catto, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs24_d1_questions.pdf

- Domain 3 Research Analysis (Yvette Hackett, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs24_d3_analysis.pdf
- Activity Model (see Case Study 24 Final Report)

25 Maria Guercio (2004), "InterPARES 2 Project - Case Study 25 Final Report: Legacoop of Bologna Web Site."
http://www.interpares.org/display_file.cfm?doc=ip2_cs25_final_report.pdf

- Case Study Proposal (Maria Guercio, 2004)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs25_proposal.pdf
- Characterization (Heather Daly, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs25_characterization.pdf
- Areas that Should Be Covered (Heather Daly, 2005)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs25_areas_2b_covered.pdf
- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs25_overview.pdf
- Diplomatic Analysis (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs25_diplomatic_analysis.pdf
- Domain 1 Research Questions (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs25_d1_questions.pdf
- Domain 3 Research Analysis (Carolyn Petrie, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_cs24_d3_analysis.pdf

General Studies

08 Mark Wolfe (2003), "InterPARES 2 Project - General Study 08 Final Report: Survey of Government Web Site Interactivity."
http://www.interpares.org/display_file.cfm?doc=ip2_gs08_final_report.pdf

- Overview (Peter Gagné, 2006)
  http://www.interpares.org/display_file.cfm?doc=ip2_gs08_overview.pdf

Domain 1: Records Creation and Maintenance

Cardin, Martin (2008), "InterPARES 2 Project - Domain 1 Task Force Report: Records Creation and Maintenance." Available in Part 2 and in electronic form at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_2_domain1_task_force.pdf.

Cardin, Martine (2006), "InterPARES 2 Project - Domain 1: Overview of Case Study Contexts," draft v1.1.
http://www.interpares.org/display_file.cfm?doc=ip2_cs_context_synthesis_v1-1.pdf

Domain 2: Authenticity, Accuracy and Reliability

Eppard, Philip (2008), "InterPARES 2 Project - Domain 2: Creator Guidelines - Making and Maintaining Digital Materials: Guidelines for Individuals." Available in Appendix 20, in electronic report form at http://www.interpares.org/display_file.cfm?doc=ip2_appendix_20.pdf, and in electronic booklet form at http://www.interpares.org/display_file.cfm?doc=ip2(pub)creator_guidelines_booklet.pdf.

Lee, Brent (2005), "InterPARES 2 Project - Domain 2: Authenticity, Accuracy and Reliability of Arts-related and Archival Literature," discussion paper. http://www.interpares.org/display_file.cfm?doc=ip2_aar_arts_lee.pdf

Roeder, John (2004), "InterPARES 2 Project - Domain 2: Authenticity, Accuracy and Reliability of Artworks: A Review of the Literature, with Some Notes about the Challenges Presented by Digital Media," discussion paper, draft version 2. http://www.interpares.org/display_file.cfm?doc=ip2_aar_arts_roeder_v2.pdf.

Roeder, John, Philip Eppard, William Underwood and Tracey P. Lauriault (2008), "InterPARES 2 Project - Domain 2 Task Force Report: Authenticity, Reliability and Accuracy of Digital Records in the Artistic, Scientific and Governmental Sectors." Available in Part 3 and in electronic form at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_3_domain2_task_force.pdf.

Domain 3: Methods of Appraisal and Preservation

Hackett, Yvette (2008), "InterPARES 2 Project - Domain 3 Task Force Report: Methods of Appraisal and Preservation." Available in Part 4 and in electronic form at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_4_domain3_task_force.pdf.

Hackett, Yvette (2008), "InterPARES 2 Project - Domain 3: Preserver Guidelines - Preserving Digital Records: Guidelines for Organizations." Available in Appendix 21, in electronic report form at http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_21.pdf, and in electronic booklet form at http://www.interpares.org/ip2/display_file.cfm?doc=ip2(pub)preserver_guidelines_booklet.pdf.

General Studies

11 Evelyn Peters McLellan (2006), "InterPARES 2 Project - General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation."
http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_english.pdf (English)
http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_french.pdf (French)

Modeling Cross-domain

Eastwood, Terry, Hans Hofman and Randy Preston (2008), "InterPARES 2 Project - Modeling Cross-domain Task Force Report: Modeling Digital Records Creation, Maintenance and Preservation." Available in Part 5 and in electronic form at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_5_modeling_task_force.pdf.

- Business-driven Recordkeeping (BDR) Model (diagrams and definitions). Available in Appendix 15 and electronic form at http://www.interpares.org/ip2/ip2_models.cfm# and http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_15.pdf.
- Chain of Preservation (COP) Model (diagrams and definitions). Available in Appendix 14 and in electronic form at http://www.interpares.org/ip2/ip2_models.cfm# and http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_14.pdf.

Preston, Randy (2004), "InterPARES 2 Project - Modeling Cross-domain: Walkthrough of the Manage Chain of Preservation Model Using Case Study 14 Data," draft report. http://www.interpares.org/display_file.cfm?doc=ip2_cs14_COP_model_walkthrough.pdf[2]

Preston, Randy (2007), "InterPARES 2 Project - Modeling Cross-domain: Crosswalk of Selected OAIS and Chain of Preservation Model Activities, Controls and Outputs," version 2. http://www.interpares.org/display_file.cfm?doc=ip2_oais_cop_model crosswalk_v2.pdf

Preston, Randy (2007), InterPARES 2 Project - Integrated Definition Function Modeling (IDEFØ): A Primer." http://www.interpares.org/display_file.cfm?doc=idef0_primer.pdf

Description Cross-domain

Gilliland, Anne et al. (2008), "InterPARES 2 Project - Description Cross-domain Task Force Report: Investigating the Roles and Requirements, Manifestations and Management of Metadata in the Creation of Reliable and Preservation of Authentic Digital Entities." Available in Part 6 and in electronic form at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_6_description_task_force.pdf.

Miller, Catherine and Eleanor Kleiber (2004), "InterPARES 2 Project - Description Cross-domain: Annotated Metadata Bibliography." http://www.interpares.org/display_file.cfm?doc=ip2(description)metadata_bibliography.pdf

Policy Cross-domain

Duranti, Luciana, Jim Suderman and Malcolm Todd (2008), "InterPARES 2 Project - Policy Cross-domain Task Force Report: Structuring the Relationship Between Records Creators and Preservers." Available in Part 7 and in electronic form at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_7_policy_task_force.pdf.

Duranti, Luciana, Jim Suderman and Malcolm Todd (2008), "InterPARES 2 Project - Policy Cross-domain: A Framework of Principles for the Development of Policies, Strategies and

[2] See general study 12 for another walkthrough of an earlier draft for the Chain of Preservation Model.

Standards for the Long-term Preservation of Digital Records." Available in Appendix 19 and in electronic booklet form at http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_19.pdf.

AUTHENTICITY, AUTHENTICATION AND THE LAW

Dalby, Seth (compiler) (2004), "InterPARES 2 Project - Policy Cross-domain: Authenticity/Authentication Definitions and Sources." http://www.interpares.org/display_file.cfm?doc=ip2(policy)authenticity_definitions-sources.pdf

Duranti, Luciana (2005), "InterPARES 2 Project - Policy Cross-domain: Authenticity and Authentication in the Law." http://www.interpares.org/display_file.cfm?doc=ip2(policy)authenticity-authentication_law.pdf

COPYRIGHT

Gutman, Susan, Luke Meagher and Adele Torrance (2006), "InterPARES 2 Project - Copyright Policy Annotated Bibliography, Draft version 4." http://www.interpares.org/display_file.cfm?doc=ip2(biblio)_copyright-annotated.pdf

LEGISLATION AND ARCHIVES STUDIES

Jim Suderman, Fiorella Foscarini and Erin Coulter (2005), "InterPARES 2 Project - Policy Cross-domain: Archives Legislation Study Report." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archives_legislation_study_report.pdf.

Australia

Kozak, Greg (compiler) (2004), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in Australia." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_AUSTRALIA.pdf

Kozak, Greg (compiler) (2005), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in Australia: Australian Evidence Act 1995." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_AUSTRALIA__EVIDENCE_ACT.pdf

Canada

Meagher, Luke (compiler) (2004), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in Canada—Provincial Government (Alberta)." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CANADA_ALBERTA.pdf

Meagher, Luke (compiler) (2004), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in Canada—Provincial Government (British Columbia)." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CANADA_BC.pdf

Meagher, Luke (compiler) (2004), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in Canada—Provincial Government (Manitoba)." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CANADA_MANITOBA.pdf

Meagher, Luke (compiler) (2004), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in Canada—Provincial Government (Ontario)." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CANADA_ONTARIO.pdf

Miller, Catherine (compiler) (2005), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in Canada—Canada Evidence Act." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CANADA_EVIDENCE_ACT.pdf

Miller, Catherine (compiler) (2004), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in Canada—Federal Government." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CANADA_FEDERAL.pdf

France

Kozak, Greg (compiler) (2004), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in France." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_FRANCE.pdf

Lamberterie, Isabelle de (2001), "InterPARES 2 Project - Policy Cross-domain: Reflexions sur l'etablissement e la conservation des actes authentiques." http://www.interpares.org/rdisplay_file.cfm?doc=ip2(delamberterie)_reflexions_etablissement_et_conservation_actes_authentiques.pdf

Hong Kong

Kozak, Greg (compiler) (2004), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in Hong Kong." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_HONG_KONG.pdf

Kozak, Greg (compiler) (2004), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in Hong Kong." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_HONG_KONG_supplement.pdf

Kozak, Greg (compiler) (2005), "InterPARES 2 Project - Policy Cross-domain: Hong Kong Evidence Ordinance and the Personal Data (Privacy) Ordinance." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_HONG_KONG_EVIDENCE_ACT.pdf

Italy

Foscarini, Fiorella (compiler) (2005), "InterPARES 2 Project - Policy Cross-domain: Authenticity and Authentication Issues in the Italian and European Union Legislation." http://www.interpares.org/display_file.cfm?doc=ip2(policy)authenticity-authentication_it-eu.pdf

Foscarini, Fiorella (2005), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in Italy (Consolidated Version)." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_ITALY_consolidated.pdf

Guercio, Maria (2004), "InterPARES 2 Project - Policy Cross-domain: The Preservation of the Electronic Records within the Italian E-Government Procedures." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_ITALY_e-gov_procedures.pdf

People's Republic of China

Xie, Sherry (2005), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in China - Report I: Provisions Having Impact on Preservation." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CHINA_no1.pdf.

Xie, Sherry (2005), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in China: Supplements to the Study of Archival Legislation in China (Report I)." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CHINA_SUPPLEMENT.pdf

Xie, Sherry (translator and compiler) (2005), "InterPARES 2 Project - Policy Cross-domain: Legislation Study - People's Republic of China (Report II): Worksright Law." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CHINA_WORKSRIGHT_LAW.pdf

Xie, Sherry (translator and compiler) (2005), "InterPARES 2 Project - Policy Cross-domain: Legislation Study - People's Republic of China (Report III): Access to Information in Chinese Legislation." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CHINA_ACCESS_TO_INFORMATION.pdf

Xie, Sherry (2005), "InterPARES 2 Project - Policy Cross-domain: Legislation Study - People's Republic of China (Report IV): The Concept of Authenticity in Chinese Legislation." http://www.interpares.org/display_file.cfm?doc=ip2(policy)authenticity_legislation_CHINA.pdf

Xie, Sherry (2005), "InterPARES 2 Project - Policy Cross-domain: Glossary for the Study of Archival Legislation in China." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CHINA_GLOSSARY.pdf

Singapore

Kozak, Greg (compiler) (2004), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in Singapore." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_SINGAPORE.pdf

Spain

Foscarini, Fiorella (2005), "InterPARES 2 Project - Policy Cross-domain: Authenticity and Authentication Issues in the Spanish Legislation." http://www.interpares.org/display_file.cfm?doc=ip2(policy)authenticity_SPAIN.pdf

United States of America

Miller, Catherine (compiler) (2004), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in the United States." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_UNITED_STATES.pdf

Miller, Catherine (compiler) (2005), "InterPARES 2 Project - Policy Cross-domain: Archival Legislation in the United States—Federal Rules of Evidence and Other Relevant Legislation." http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_UNITED_STATES_EVIDENCE_ACT.pdf

MORAL RIGHTS

Taitano, Melissa, and Sharon Farb (2005), "InterPARES 2 Project - Policy Cross-domain: International Moral Rights." http://www.interpares.org/display_file.cfm?doc=ip2(policy)moral_rights.pdf

POLICIES, GUIDELINES AND STANDARDS

Morrison, Jane (2004), "InterPARES 2 Project - Policy Cross-domain: A Survey of Policies, Strategies, Guidelines, Standards and Best Practices of Arts Organizations in Canada." http://www.interpares.org/display_file.cfm?doc=ip2(policy)arts_CANADA.pdf

PRIVACY

Maxwell, Terry (2004), "InterPARES 2 Project - Policy Cross-domain: United States Privacy Research." http://www.interpares.org/display_file.cfm?doc=ip2(policy)privacy_USA.pdf

Todd, Malcolm (2005), "InterPARES 2 Project - Policy Cross-domain Information Policy: Privacy Report." http://www.interpares.org/display_file.cfm?doc=ip2(policy)privacy_report.pdf

Iacovino, Livia (2004), "InterPARES 2 Project - Policy Cross-domain: Provisions in European Union National Privacy Laws: Historical/archival/research Exemptions by Country," working version 1.0. http://www.interpares.org/display_file.cfm?doc=ip2(policy)privacy_EU.pdf

Terminology Cross-domain

Tennis, Joseph T. and Randy Preston (2008), "InterPARES 2 Project - Terminology Cross-domain Task Force Report: Terminological Instruments." Available in Part 8 and in electronic form at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_8_terminology_task_force.pdf.

- InterPARES 2 Project Dictionary. Available in electronic report form at http://www.interpares.org/ip2/ip2_term_pdf.cfm?pdf=dictionary and in the online Terminology Database at http://www.interpares.org/ip2/ip2_terminology_db.cfm.
- InterPARES 2 Project Glossary. Available in this book, in electronic report form at http://www.interpares.org/display_file.cfm?doc=ip2_book_glossary.pdf, and in the online Terminology Database at http://www.interpares.org/ip2/ip2_terminology_db.cfm.
- InterPARES 2 Project Ontologies. Available in Appendix 22 and in electronic form at http://www.interpares.org/ip2/display_file.cfm?doc=ip2_ontology.pdf.
- InterPARES 2 Terminology Database. Available only online at http://www.interpares.org/ip2/ip2_terminology_db.cfm

BIBLIOGRAPHY

Bibliography[1]

Abrams, Stephen L. (2005), "Establishing a Global Digital Format Registry," *Library Trends* 54(1): 125–143. http://muse.jhu.edu/demo/library_trends/v054/54.1abrams.pdf

American Library Association, "Digital Rights Issues," Washington Policy Office. http://www.ala.org/ala/washoff/WOissues/copyrightb/digitalrights/DRMissues.pdf

An, Xiaomi (2003), "An Integrated Approach to Records Management," *Information Management Journal* July/August: 24–30.

Appraisal Task Force (2001), "Appendix 4: A Model of the Selection Function," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 239–252. Online reprint available at http://www.interpares.org/display_file.cfm?doc=ip1_aptf_model.pdf

Archaeological Institute of America, *AIA Code of Professional Standards*. http://www.archaeological.org/pdfs/AIA_Code_of_Professional_StandardsA5S.pdf

Arnheim, Rudolf (1993), "The Two Authenticities of the Photographic Media," *Journal of Aesthetics and Art Criticism* 51(4): 537–540.

Authenticity Task Force (2000), "Appendix 1: Template for Analysis," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 192–203. Online reprint available at http://www.interpares.org/book/interpares_book_j_app01.pdf

Authenticity Task Force (2002), "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records" in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 204–219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf

Avedon, Richard, Foreword to *In the American West 1979-1984* (New York: Harry N. Abrams, 1985). http://www.richardavedon.com/#mi=1&pt=0&pi=11019&p=-1&at=-1

Baca, Murtha (ed.), *Introduction to Metadata: Pathways to Digital Information* (Los Angeles: Getty Information Institute, 1998).

Bachimont, Bruno, Jean-François Blanchette, Andrew Gerzso, Anne Swetland, Oliver Lescurieux, Pierre Morizet-Mahoudeaux, Nicolas Nodin and Jill Teasley. "Preserving Interactive Digital Music: A Report on the MUSTICA Initiative," in *Proceedings of the Third International Conference on Web Delivering of Music (WEDELMUSIC 2003)*, 15–17 September 2003, Leeds, UK (Washington, D.C.: IEEE Computer Society Press, 2003). http://polaris.gseis.ucla.edu/blanchette/papers/wedelmusic.pdf

Beagrie, Neil and Daniel Greenstein, *A Strategic Policy Framework for Creating and Preserving Digital Collections. JISC/NPO Studies on the Preservation of Electronic Materials*, eLib Supporting Study P3, version 4 (London: South Bank University, Library Information Technology Centre, 1998). http://www.ukoln.ac.uk/services/papers/bl/framework/framework.html Version 5.0 (2001) http://ahds.ac.uk/strategic.pdf

Berman, Fran, Geoffrey Fox and Tony Hey (eds.), *Grid Computing: Making the Global Infrastructure a Reality* (Chichester, West Sussex: John Wiley & Sons, 2003).

[1] Excludes InterPARES 2 Project case study, general study and research unit documents (see Appendix 23).

Berman, Helen M., John Westbrook, Zukang Feng, Gary Gilliland, T. N. Bhat, Helge Weissig, Ilya N. Shindyalov and Philip E. Bourne (2000), "The Protein Data Bank," *Nucleic Acids Research* 28(1): 235–242. http://nar.oxfordjournals.org/cgi/content/full/28/1/235

Bernardini, Nicola and Alvise Vidolin (2005), "Sustainable Live Electroacoustic Music," *eContact!* 8(3). http://cec.concordia.ca/econtact/8_3/bernardini_vidolin.html

Bos, J. J. (1996), "Digital Signatures and the Electronic Health Records: Providing Legal and Security Guarantees," *International Journal of Bio-Medical Computing* 42(1-2): 157–163.

Bradshaw, David (1998), EBU-SMPTE Task Force on Metadata. "Digital Workflow Through Production and Documentation." A Seminar Organised by the Documentation Commission of FIAT/IFTA, BBC Conference Centre, London, May 7-8, 1998.

Brown, David L., Grace Welch and Christine Cullingworth (2005), "Archiving, Management and Preservation of Geospatial Data: Summary Report and Recommendations," GeoConnections Policy Advisory Node: Working Group on Archiving and Preserving Geospatial Data. http://www.geoconnections.org/publications/policyDocs/keyDocs/geospatial_data_mgt_summary_report_20050208_E.pdf

Buskirk, Martha, *The Contingent Object of Contemporary Art* (Cambridge, MA: MIT Press, 2005).

Canadian Archaeological Association, *Principles of Ethical Conduct.* http://www.canadianarchaeology.com/conduct.lasso

Center for Research Libraries, "Auditing and Certification of Digital Archives." http://www.crl.edu/content.asp?l1=13&l2=58&l3=162

Center for Technology in Government (1998), "Functional Requirements to Ensure the Creation, Maintenance and Preservation of Electronic Records: Models for Action Project: Developing Practical Approaches to Electronic Records Management and Preservation," CTG.MFA - 004, Center for Technology in Government, University at Albany/SUNY and New York State Archives. http://www.ctg.albany.edu/publications/reports/functional/functional.pdf

Chadabe, Joel (2001), "Preserving Performances of Electronic Music," *Journal of New Music Research* 30(4): 303–305.

China, *Electronic Signatures Law of the People's Republic of China*.

Chilvers, Alison and John Feather (1998), "The Management of Digital Data: A Metadata Approach," *Electronic Library* 16(6): 365–372.

Clarke, Derek G. and David M. Clark, "Chapter 2: Lineage," in *Elements of Data Quality*, Stephen C. Guptill and Joel L. Morrison, eds. (Oxford: Elsevier Science, 1995), 13–30.

Cloonan, Michèle V. and Shelby Sanett (2002), "Where We Are Now: Obliquity and Squint?" *The American Archivist* 65(1): 70–106.

Cloonan, Michèle, V. and Shelby Sanett (2001), "Survey of Preservation Practices and Plans," draft appendix to the InterPARES 1 Preservation Task Force Final Report. http://www.interpares.org/display_file.cfm?doc=ip1_survey_of_preservation_practices-plans.pdf

CODATA Working Group on Archiving Scientific Data, The Committee on Data for Science and Technology (CODATA) of the International Council for Science. http://www.nrf.ac.za/codata/

Cook, Terry (1995), "It's Ten O'Clock, Do You Know Where Your Data Are?" *Technology Review* 98: 48–53.

Cumming, Kate, "Purposeful Data: The Roles and Purposes of Recordkeeping Metadata" (Ph.D. dissertation, Monash University, 2005).

Davies, David, *Art as Performance* (Oxford: Blackwell, 2004).

Davies, Stephen (1991), "The Ontology of Musical Works and the Authenticity of their Performance." *Nous* 25: 21–41. Reprinted in *Themes in the Philosophy of Music* (New York: Oxford University Press, 2003), 60–77.

Dingwall, Glenn, Richard Marciano, Reagan Moore and Evelyn Peters McLellan (2008), "From Data to Records: Preserving the Geographic Information System of the City of Vancouver," *Archivaria* (Fall): 181–198.

Duff, Wendy M. "The Influence of Warrant on the Acceptance and Credibility of the Functional Requirements for Recordkeeping" (Ph.D. dissertation, University of Pittsburgh,1996).

Duff, Wendy M. (1997), "Compiling Warrant in Support of the Functional Requirements for Recordkeeping," *Bulletin of the American Society for Information Science* 23(5): 12–13. http://doi.wiley.com/10.1002/bult.60

Duff, Wendy M. (1997), "Warrant and the Definition of Electronic Records: Questions Arising from the Pittsburgh Project," *Archives and Museum Informatics* 11(3-4): 223–231. http://tort.library.utoronto.ca:8080/bitstream/1778/4315/1/Warrent.pdf

Duranti, Luciana, *Diplomatics: New Uses for an Old Science* (Lanham, Maryland and London: The Scarecrow Press in association with the Society of American Archivists and the Association of Canadian Archivists, 1998).

Duranti, Luciana (2001), "International Research on Permanent Authentic Records in Electronic Systems (InterPARES): Experiential, Interactive and Dynamic Records," SSHRC MCRI InterPARES 2 Project Proposal, 412-2001. http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf

Duranti, Luciana (2004), "InterPARES 2 Project Midterm Report to the Social Sciences and Humanities Research Council." MCRI Grant No. 412-2001-1003. May 25, 2004.

Duranti, Luciana (2005), "InterPARES 2 Project - The Concept of the Records Life Cycle," PowerPoint presentation. (unpublished)

Duranti, Luciana (ed.), *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* (San Miniato, Italy: Archilab, 2005). Online reprint available at http://www.interpares.org/book/index.cfm

Duranti, Luciana et al., "Part Four – An Intellectual Framework for Policies, Strategies, and Standards: Strategy Task Force Report," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 117–125. Online reprint available at http://www.interpares.org/book/interpares_book_g_part4.pdf

Duranti, Luciana and Heather MacNeil (1996), "The Protection of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project," *Archivaria* 42 (Fall): 46–67. Online reprint available at http://journals.sfu.ca/archivar/index.php/archivaria/article/view/12153/13158

Duranti, Luciana and Kenneth Thibodeau (2006), "The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES," *Archival Science* 6(1): 13–68.

Duranti, Luciana, Terry Eastwood and Heather MacNeil, *Preservation of the Integrity of Electronic Records* (Dordrecht: Kluwer Academic Publishers, 2002). Online reprint available at http://www.interpares.org/UBCProject/index.htm

Eastwood, Terry et al., "Part Two – Choosing to Preserve: The Selection of Electronic Records: Appraisal Task Force Report," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 67–98. Online reprint available at http://www.interpares.org/book/interpares_book_e_part2.pdf

European Task Force Permanent Access (2005), "Permanent Access to the Records of Science: Strategic Action Programme." http://www.alliancepermanentaccess.eu/Proposal%20Research%20and%20Development.doc

Evans, Joanne and Lori Lindberg, "Describing and Analyzing the Recordkeeping Capabilities of Metadata Sets," in *DC-2004: Proceedings of the International Conference on Dublin Core and Metadata Applications, October 11-14 2004, Shanghai, China* (Shanghai, China: Shanghai Scientific and Technological Literature Publishing House, 2004), 75–80. Online reprint available at http://www.dublincore.go.kr/dcpapers/pdf/2004/Paper_27.pdf

Evans, Joanne and Nadav Rouche (2004), "Utilizing Systems Development Methods in Archival Systems Research: Building a Metadata Schema Registry," *Archival Science* 4(3-4): 315–334.

Faundeen, John L. (2003), "The Challenge of Archiving and Preserving Remotely Sensed Data," *Data Science Journal* 2: 159–163. http://www.jstage.jst.go.jp/article/dsj/2/0/159/_pdf

Flynn, Sarah J. A. (2001), "The Records Continuum Model in Context and its Implications for Archival Practice," *Journal of the Society of Archivists* 22(1): 79–93.

Gavrel, Katherine, *Conceptual Problems Posed by Electronic Records: A RAMP Study*, PGI-90/WS/12 (Paris: UNESCO, 1990). http://unesdoc.unesco.org/images/0008/000866/086689eo.pdf

Gerzso, Andrew (2005), "Performance Handbook: *Anthèmes 2* [by Pierre Boulez]" (Paris: IRCAM). http://mustica.ircam.fr/mustica_1.2.0/rendu/pdf/output/Anthemes_2.pdf

Gilliland, Anne (2005), "Discussion Paper on the Nature and Role of Metadata in the Creation of Reliable and the Preservation of Authentic Records in Electronic Systems," paper presented at InterPARES 2 Project Plenary Workshop 13, 20-24 February 2005, Vancouver, BC, Canada. (unpublished)

Gilliland, Anne J., Nadav Rouche, Joanne Evans and Lori Lindberg (2005), "Towards a Twenty-First Century Metadata Infrastructure Supporting the Creation, Preservation and Use of Trustworthy Records: Developing the InterPARES 2 Metadata Schema Registry," *Archival Science* 4(1): 43–78.

Gilliland-Swetland, Anne J. (1999), "Metadata and the Management of Authentic Electronic Records," paper presented at the EDUCAUSE Conference, October 1999, Long Beach, CA, USA.

Gilliland-Swetland, Anne J. and Sue McKemmish, "A Metadata Schema Registry for the Registration and Analysis of Recordkeeping and Preservation Metadata," in *Proceedings of the Second IS&T Archiving Conference, April 26-29, 2005, Washington, D.C.* (Springfield, VA: Society for Imaging Science and Technology, 2005), 109–112.

Glick, Kevin (2006), "Electronic Records Preservation Strategies," (unpublished).

Goodman, Nelson, "Authenticity," in *The Dictionary of Art*, Jane Turner, ed. (New York: Grove, 1996), 834–835.

Goodrich, Michael T., Roberto Tamassia, Nikos Triandopoulos and Robert Cohen (2001), "Authenticated Data Structures for Graph and Geometric Searching." Technical report, Center for Geometric Computing, Brown University. http://www.cs.brown.edu/cgc/stms/papers/authDataStr.pdf

Gourad, Khalid, "Geographic Information Systems in Archaeology: A Survey" (unpublished Masters thesis, Hunter College of the City University of New York, Department of Anthropology, 1999), 75 pp. http://khalid.gourad.com/thesis/

Greenwood, Mark, Carole Goble, Robert Stevens, Jun Zhao, Matthew Addis, Darren Marvin, Luc Moreau and Tom Oinn (2003), "Provenance of e-Science Experiments: Experience from Bioinformatics," in *Proceedings UK e-Science All Hands Meeting 2003*, Simon J. Cox, ed. http://www.nesc.ac.uk/events/ahm2003/AHMCD/pdf/047.pdf

Grout, Catherine, Phil Purdy and Janine Rymer (2000), *Creating Digital Resources for the Visual Arts: Standards and Good Practice,* Visual Arts Data Service and Technical Advisory Service for Images (UK). http://vads.ahds.ac.uk/guides/creating_guide/contents.html

Guptill, Stephen G. and Joel L. Morrison (eds.), *Elements of Spatial Data Quality* (Oxford: Elsevier Science, 1995).

Hanrahan, Robert P. (1995), "The IDEF Process Modeling Methodology." http://www.stsc.hill.af.mil/crosstalk/1995/06/IDEF.asp

Harries, Stephen (1999), "Capturing and Managing Electronic Records from Websites and Intranets in the Government Environment," paper presented at the DLM-Forum on Electronic Records of the European Communities, 18-19 October 1999, Brussels, Belgium. http://ec.europa.eu/archives/ISPO/dlm/fulltext/full_harr_en.htm

Harvey, Ross (2000), "An Amnesiac Society? Keeping Digital Data for Use in the Future." Paper presented at the LIANZA 2000 Conference, New Zealand, 15-18 October 2000.

Hass, Joan K., Helen W. Samuels and Barbara T. Simmons, *Appraising the Records of Modern Science and Technology: A Guide* (Cambridge, MA: Massachusetts Institute of Technology, 1985).

Hass, Joan K., Helen W. Samuels and Barbara T. Simmons (1986), "The MIT Appraisal Project and its Broader Implications," *American Archivist* 49(3): 310–314.

Heslop, Helen, Simon Davis and Andrew Wilson (2002), "An Approach to the Preservation of Digital Records," National Archives of Australia. http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm2-888.pdf

Hodge, Sharon H. (1997), "Satellite Data and Environmental Law: Technology Ripe for Litigation Application," *Pace Environmental Law Review* 14: 691–718.

Hurley, Chris (1995), "Ambient Functions: Abandoned Children to Zoos," *Archivaria* 40 (Fall): 21–39. Online reprint available at http://journals.sfu.ca/archivar/index.php/archivaria/article/viewFile/12095/13080

Iacovino, Livia and Malcolm Todd (2007), "The Long-Term Preservation of Identifiable Personal Data: A Comparative Archival Perspective on Privacy Regulatory Models in the European Union, Australia, Canada and the United States," *Archival Science* 7(1): 107–127.

International Organization for Standardization, ISO 14721: 2003 - Space data and information transfer systems—Open archival information system—Reference model.

International Organization for Standardization, ISO 15489-1:2001 - Information and documentation—Records management—Part 1: General.

International Organization for Standardization, ISO 15836-1:2003 - Information and documentation—The Dublin Core metadata element set.

International Organization for Standardization, ISO 19005-1:2005 - Document management—Electronic document file format for long-term preservation—Part 1: Use of PDF 1.4 (PDF/A-1).

International Organization for Standardization, ISO 19113:2002 - Geographic Information—Quality Principles.

International Organization for Standardization, ISO 19114:2003 - Geographic Information —Quality evaluation procedures.

International Organization for Standardization, ISO 23081-1:2006 - Information and documentation—Records management processes—Metadata for records—Part 1: Principles.

International Organization for Standardization, International Electrotechnical Commission, ISO/IEC 11179: Information Technology—Metadata registries (MDR). http://metadata-stds.org/11179/index.html

InterPARES 2 Project, "Milestones Report," October 2002. http://www.interpares.org/display_file.cfm?doc=ip2_milestones(200210).pdf

InterPARES 2 Project, "Overview of InterPARES 2 Intellectual Framework." http://www.interpares.org/display_file.cfm?doc=ip2_overview_of_intellectual_framework(20030311).pdf

InterPARES 2 Project, "Proceedings of Workshop #2," National Archives and Records Administration, College Park, Maryland, USA, 20-22 June 2002. Summary version at http://www.interpares.org/display_file.cfm?doc=ip2_wk02_summary.pdf

InterPARES 2 Project, "Summary of Proceeding of Workshop #3," Crowne Plaza Beverly Hills, Los Angeles, CA, USA, 17-21 September 2002. http://www.interpares.org/display_file.cfm?doc=ip2_wk03_summary.pdf

Italy, DPR 445/2000. http://www.parlamento.it/parlam/leggi/deleghe/00443dla.htm

Jones, Caitlin (2004), "Does Hardware Dictate Meaning? Three Variable Media Conservation Case Studies," *horizon*^{0} 18.2. http://www.horizonzero.ca/textsite/ghost.php?is=18&file=6&tlang=0

Jøsang, Audun and Mary Anne Patton,"User Interface Requirements for Authentication of Communication," in *Proceedings of the Fourth Australian User interface Conference on User interfaces 2003 - Volume 18*, R. Biddle and B. Thomas, eds. ACM International Conference Proceeding Series, vol. 36 (Darlinghurst, Australia: Australian Computer Society, 2003), 75–80. http://portal.acm.org/ft_gateway.cfm?id=820105&type=pdf&coll=&dl=&CFID=15151515&CFTOKEN=6184618

Kennedy, Jay and Cherryl Schauder, *Records Management: A Guide to Corporate Recordkeeping*, 2nd edition (South Melbourne: Longman, 1998).

Kintigh, Keith W. (1996), "Principle No. 1: Stewardship," in *SAA Principles of Archaeological Ethics*. http://www.saa.org/Publications/SAAbulletin/14-3/SAA9.html

Kivy, Peter, *Authenticities: Philosophical Reflections on Musical Performance* (Ithaca and London: Cornell University Press, 1995).

Lauriault, Tracey P. and Barbara Craig (2006), "Do Data Access Portals, Repositories, and Catalogues, Preserve or Archive Geospatial and Science Data?" paper presented at the GeoTech 2006 Conference, 18-21 June 2006, Ottawa, ON, Canada. http://www.interpares.org/display_file.cfm?doc=ip2_gs10_geotec2006.pdf

Lauriault, Tracey P., Barbara L. Craig, D. R. Fraser Taylor and Peter L. Pulsifer (2007), "Today's Data are Part of Tomorrow's Research: Archival Issues in the Sciences," *Archivaria* 64 (Fall): 123–179.

Lee, Brent (2001), "The Growing Complexity of Music Preservation," in *Preserving Authentic Electronic Records: Preliminary Research Findings, an International Symposium held February 17, 2001 at the University of British Columbia, Vancouver, Canada,* Luigi Sarno, ed. , 32–40. Online reprint available at http://www.interpares.org/display_file.cfm?doc=ip1_symposium_2001.pdf.

Lee, Brent and John Roeder, "The Challenge of Digital Musical Archives: Research within the InterPARES Project," in *Digital Resources for the Humanities 2001-2002*, J. Anderson, A. Dunning and M. Fraser, eds. (London: Office for Humanities Communication [King's College], 2003), 171–180.

Lessig, Lawrence, "Editorial: The Coming of Copyright Perpetuity," *New York Times*, January 16, 2003, p. A28.

Lessig, Lawrence, *Free Culture* (The Penguin Press, 2004).

Levinson, Jerrold, "Autographic and Allographic Art Revisited," in *Music, Art, and Metaphysics: Essays in Philosophical Aesthetics* (Ithaca: Cornell University Press, 1990), 89–106.

Library of Congress, *National Digital Information Infrastructure and Preservation Program (NDIIPP), Digital Preservation*. http://www.digitalpreservation.gov/index.html

Lindberg, Lori, Monique Leahey-Sugimoto, Nadav Rouche and Holly Wang, "MADRAS: A Metadata and Archival Description Registration and Analysis System for the Analysis of the Recordkeeping Capabilities of Metadata Sets," in *Proceedings of the Third IS&T Archiving Conference* (Springfield, VA: Society for Imaging Science and Technology, 2006), 216–218.

Liu, Alan, David Durand, Nick Montfort, Merrilee Proffitt, Liam R. E. Quin, Jean-Hughes Réty and Noah Wardrip-Fruin. "Born-Again Bits: A Framework for Migrating Electronic Literature." Version 1.1, August 5, 2005. The Electronic Literature Organization. http://www.eliterature.org/pad/bab.html

MacNeil, Heather and Bonnie Mak (2007), "Constructions of Authenticity," *Library Trends: Recent Trends in Cultural Heritage Preservation* 56(1): 26–52.

MacNeil, Heather, et al., "Part One – Establishing and Maintaining Trust in Electronic Records: Authenticity Task Force Report," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 19–65. Online reprint available at http://www.interpares.org/book/interpares_book_d_part1.pdf

Marshall, Peter (2000), "Life Cycle Versus Continuum: What is the Difference?" *Informaa Quarterly* 16(2): 20–25.

Martel, Chip, Glen Nuckolls, Michael Gertz, Prem Devanbu, April Kwong and Stuart Stubblebine (2001), "A General Model for Authentic Data Publication." Available at http://www.cs.ucdavis.edu/~devanbu/files/model-paper.pdf

McCargar, Victoria (2005), "Following the Trail of the Disappearing Data," *The Seybold Report* 4(21): 7–14.

McCargar, Victoria (2006), "The Heart of Darkness: A Foray into Aging JPEGs," *The Seybold Report* 5(22): 9–12.

McCargar, Victoria (2006), "You Can Kiss Your Assets Goodbye: The State of News Archives at the Dawn of the Digital Era," paper presented at the Special Libraries Association Annual Conference 2006, June 10-15, 2006, Baltimore, MD, USA. http://www.ibiblio.org/slanews/conferences/sla2006/presentations/assets.pdf (text), http://www.ibiblio.org/slanews/conferences/sla2006/audio/vicky.mp3 (audio)

McCargar, Victoria and Peter F. Johnson (2006), "Comments to Section 108 Study Group: News Archives," submitted April 28, 2006. http://www.loc.gov/section108/docs/McCargar.pdf

McDonald, John (2000), "Archives and Current Records: Towards a Set of Guiding Principles," Committee on Electronic and Other Current Records, International Council on Archives. August 2000. http://www.ica.org/sites/default/files/principles_eng.html

McGraw-Hill, *Dictionary of Scientific and Technical Terms*, sixth edition (New York: McGraw-Hill, 2003).

McKemmish, Sue (1997), "Yesterday, Today and Tomorrow: A Continuum Responsibility," in *Proceedings of the Records Management Association of Australia 14th National Convention*, RMAA Perth, 15-17 September 1997.

McKemmish, Sue, Glenda Ackland, Nigel Ward and Barbara Reed (1999), "Describing Records in Context in the Continuum: The Australian Recordkeeping Metadata Schema," *Archivaria* 48 (Fall): 3–37. http://journals.sfu.ca/archivar/index.php/archivaria/article/view/12715/13890

McKemmish, Sue, Michael Piggott, Barbara Reed and Frank Upward (eds.), *Archives: Recordkeeping in Society* (Wagga Wagga, Australia: Centre for Information Studies, 2005).

Meijer, Albert (2001), "Electronic Records Management and Public Accountability: Beyond an Instrumental Approach," *The Information Society* 17(4): 259–270.

Miller, Joanne, "Chapter 2: The Cost of Creating Digital Images and Metadata by Museums," in *The Cost of Digital Image Distribution: The Social and Economic Implications of the Production, Distribution, and Usage of Image Data*, Howard Besser and Robert Yamashita, eds.(Berkeley: UCB School of Information Management & Systems, 1998). Reprinted in *Visual Resources* 14(4): 399–411. http://www2.sims.berkeley.edu/research/projects/mellon/mellon_pdfs/f2-muse-0202.pdf

Minitski, A., A. Mogilner, C. MacKnight and K. Rockwood (2003), "Data Integration and Knowledge Discovery in Biomedical Databases. Reliable Information from Unreliable Sources," *Data Science Journal* 2: 25–34. http://journals.eecs.qub.ac.uk/codata/journal/contents/2_03/2_03pdfs/DS131.pdf

Minnesota Historical Society, State Archives Department (2002), *Trustworthy Information Systems Handbook*. Version 4. http://www.mnhs.org/preserve/records/tis/tableofcontents.html

Montfort, Nick and Noah Wardrip-Fruin (2004), "Acid-Free Bits: Recommendations for Long-Lasting Electronic Literature." Version 1.0, June 14, 2004. The Electronic Literature Organization. http://www.eliterature.org/pad/afb.html

Moore, Allan (2002), "Authenticity as authentication," *Popular Music* 21(2): 209–223. http://journals.cambridge.org/article_S0261143002002131

Moore, Reagan, Chaitan Baru, Arcot Rajasekar, Bertram Ludaescher, Richard Marciano, Michael Wan, Wayne Schroeder and Amarnath Gupta (2000), "Collection-Based Persistent Digital Archives - Part 1," *D-Lib Magazine* 6(3). http://www.dlib.org/dlib/march00/moore/03moore-pt1.html

Moore, Reagan, Chaitan Baru, Arcot Rajasekar, Bertram Ludaescher, Richard Marciano, Michael Wan, Wayne Schroeder and Amarnath Gupta (2000), "Collection-Based Persistent Digital Archives - Part 2," *D-Lib Magazine* 6(4). http://www.dlib.org/dlib/april00/moore/04moore-pt2.html

Myburgh, Susan (1998), "Metadata and its Meaning for Records Managers," *Records & Information Management Report* 14(5).

National Aeronautics and Space Administration, "Chapter 12. Object Description Language Specification and Usage," in *Planetary Data System Standards Reference*, version 3.7, March 20, 2006. JPL D-7669, Part 2 (Pasadena, CA: Jet Propulsion Laboratory, California Institute of Technology, 2006). http://pds.nasa.gov/documents/sr/

National Archives of Australia (2004), *Digital Recordkeeping: Guidelines for Creating, Managing and Preserving Digital Records*. http://www.naa.gov.au/Images/Digital-recordkeeping-guidelines_tcm2-920.pdf

National Archives of Canada (2001), *Guidelines for Records Created Under a Public Key Infrastructure Using Encryption and Digital Signatures*. http://www.lac-bac.gc.ca/information-management/002/007002-3015-e.html

National Electronic Commerce Coordinating Council, *Creating and Maintaining Proper Systems for Electronic Record Keeping* (Albany: SUNY Albany Center for Technology in Government, 2002). Report presented at the NECCC Annual Conference, December 4-6, 2002, New York, NY, USA. http://www.ctg.albany.edu/publications/reports/proper_systems/proper_systems.pdf

National Research Council, Chemical Sciences Roundtable, *Impact of Advances in Computing and Communications Technologies on Chemical Science and Technology: Report of a Workshop* (Washington, D.C.: National Academy Press, 1999). http://www.nap.edu/catalog.php?record_id=9591#toc

National Research Council, Committee on Data Management and Computation, *Issues and Recommendations Associated with Distributed Computation and Data Management Systems for the Space Sciences*, report no. NAS 1.26183026 (Washington, D.C.: National Academy Press, 1986). http://hdl.handle.net/2060/19880017724.

National Research Council, Commission on Physical Sciences, Mathematics and Applications, *Preserving Scientific Data on Our Physical Universe: A New Strategy for Archiving the Nation's Scientific Information Resources* (Washington, D.C.: National Academy Press, 1995). http://www.nap.edu/catalog.php?record_id=4871#toc

National Research Council, Committee on Issues in the Transborder Flow of Scientific Data, *Bits of Power: Issues in Global Access to Scientific Data* (Washington, D.C.: National Academy Press, 1997). http://www.nap.edu/catalog.php?record_id=5504#toc

National Research Council, Steering Committee for the Study on the Long-term Retention of Selected Scientific and Technical Records of the Federal Government, *Study on the Long-term Retention of Selected Scientific and Technical Records of the Federal Government: Working Papers* (Washington, D.C.: National Academy Press, 1995). http://books.nap.edu/catalog.php?record_id=9478#toc

National Science Foundation (2005), *Report of the National Science Board: Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century,* NSB-05-40, September 2005. http://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf

Navab, Aphrodite Désirée (2003), Review of *Transforming Images: How Photography Complicates the Picture*, by Barbara E. Savedoff. *Journal of Aesthetic Education* 37(2): 114–121.

New York State Office for Technology (2002), "E-Records Guidelines Ensuring the Security, Authenticity, Integrity, and Accessibility of Electronic Records," Part 4 of *Electronic Signature and Records Act (ESRA) Guidelines*. http://www.oft.state.ny.us/arcPolicy/policy/ESRAGuidelines4.htm

New York State Office for Technology (2002), "Glossary," Part 5 of *Electronic Signature and Records Act (ESRA) Guidelines*. http://www.oft.state.ny.us/arcPolicy/policy/ESRAGuidelines5.htm

New York Times Co., Inc., et al. v. Tasini et al. (00-201) 533 U.S. 483 (2001) 206 F.3d 161, affirmed. http://supct.law.cornell.edu/supct/html/00-201.ZS.html

Online Computer Library Center, Center for Research Libraries (2007), "Trustworthy Repositories Audit & Certification: Criteria and Checklist," v. 1.0, February 2007. http://www.crl.edu/PDF/trac.pdf

Pearce-Moses, Richard, *A Glossary of Archival and Records Terminology* (Chicago: Society of American Archivists, 2005). http://www.archivists.org/glossary/

Pilone, Dan, *UML 2.0 Pocket Reference* (O'Reilly Media, 2006).

Preservation Task Force (2002), "Appendix 5: A Model of the Preservation Function," version 6.0, in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 253–292. Online reprint available at http://www.interpares.org/display_file.cfm?doc=ip1_ptf_model.pdf

Records Continuum Research Group (1998), "Create Once, Use Many Times - The Clever Use of Metadata in eGovernment and eBusiness Processes in Networked Environments." http://www.sims.monash.edu.au/research/rcrg/research/crm/

Reilly, Bernard F. Jr., "Knowledge Biodiversity: The Perilous Economic of World News Heritage Materials," in *Proceedings of the ACRL Twelfth National Conference, April 7-10, 2005, Minneapolis, MN* (Minneapolis, MN: Association of College and Research Libraries, 2005), 238–243. http://www.ala.org/ala/acrl/acrlevents/reilly05.pdf

Rohrbaugh, Guy, "Ontology of Art," in *The Routledge Companion to Aesthetics.* 2nd ed. B. Gaut and D. M. Lopes, eds. (New York: Routledge, 2005), 241–254.

Ross, Seamus (1999), "Report on "Creation, preservation and access to electronic information: economical and functional aspects" - Maintaining the value of intellectual capital," paper presented at the DLM-Forum on Electronic Records of the European Communities, 18-19 October 1999, Brussels, Belgium. http://ec.europa.eu/archives/ISPO/dlm/fulltext/full_ross_en.htm

San Diego Supercomputer Center (1999), "Methodologies for the Long-Term Preservation of and Access to Software-Dependent Electronic Records," NHPRC Proposal, June 1, 1999. http://www.sdsc.edu/NARA/Publications/nhprc_latest.pdf (full text), http://www.sdsc.edu/NARA/Publications/nhprc_summary.pdf (summary)

Simmhan, Yogesh L., Beth Plale and Dennis Gannon (2005), "A Survey of Data Provenance in e-Science," *Special Interest Group on Management of Data (SIGMOD) Record* 34(3): 31–36. http://www.sigmod.org/sigmod/record/issues/0509/p31-special-sw-section-5.pdf

Sirven, Xavier (2004), "Authenticité et accessibilité des archives électroniques - MUSTICA, Le cas de la création musicale numérique," Technical Report, Université Technologique de Compiègne. http://polaris.gseis.ucla.edu/blanchette/papers/RapportSirven.pdf

Smith, Barry (ed.), *Guide to Good Practice Creating Digital Performance Resources*, AHDS Performing Arts Data Service. http://ahds.ac.uk/creating/guides/performing-resources/contents.htm

Spoo, Robert (1998), "Copyright Protectionism and Its Discontents: The Case of James Joyce's *Ulysses* in America," *Yale Law Journal* 108(3): 633–667.
Standards Australia, *Australian Standard 4390, Records Management* (Homebush, New South Wales: Standards Australia, 1996).
Strong, David F. and Peter B. Leach, *National Consultation on Access to Scientific Data: Final Report* (Ottawa: Canadian Institute for Scientific and Technical Information, National Research Council Canada, 2005). http://ncasrd-cnadrs.scitech.gc.ca/NCASRDReport_e.pdf
Tennis, Joseph T. (2008), "Metadata in the Chain of Preservation Model," *Archivaria* (in press)
Texas Department of Information Resources (2004), *Architecture Components for the Enterprise*. *Data and E-Records Management Domain*. http://www.dir.state.tx.us/ace/documents/phase1toc.htm
Thibodeau, Kenneth (1995), "Preserving Scientific Data on Our Physical Universe," *IASSIST Quaterly* 19(4): 26–29. http://iassistdata.org/publications/iq/iq19/iqvol194thibodeau.pdf
Thibodeau, Kenneth (2001), "Building the Archives of the Future," *D-Lib Magazine* 7(2). http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html
Thibodeau, Kenneth et al., "Part Three – Trusting to Time: Preserving Authentic Records in the Long Term: Preservation Task Force Report," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 99–116. Online reprint available at http://www.interpares.org/book/interpares_book_f_part3.pdf
Thompson, Jerry L. (2002), "Truth and Photography," *Yale Review* 90(1): 25–53.
Tibbo, Helen R. (2003), "Primarily History in America: How U.S. Historians Search for Primary Materials at the Dawn of the Digital Age," *American Archivist* 66(1): 9–50.
Todd, Malcolm (2006), "Power, Identity, Integrity, Authenticity, and the Archives: A Comparative Study of the Application of Archival Methodologies to Contemporary Privacy," *Archivaria* 61 (Spring): 181–214.
Underwood, E. William (2002), "A Formal Method for Analyzing the Authenticity Properties of Procedures for Preserving Digital Records," in *Proceedings of the 2002 International Conference on Digital Archive Technologies (ICDAT2002),* December 19-20, 2002, Academia Sinica, Taipei, Taiwan, 53–64. http://perpos.gtri.gatech.edu/publications/ICDAT2002.pdf#page=1
United States Congress, *E-Government Act of 2002*. 107th Cong., 2d sess., 2002. H.R.2458.ENR.
United States House of Representatives, *Federal Rules of Evidence*, Article X. Contents of Writings, Recordings, and Photographs: Rule 1001. Definitions, Committee on the Judiciary, Committee Print No. 8 (December 31, 2004). http://judiciary.house.gov/media/pdfs/printers/108th/evid2004.pdf
United States Secretary of Commerce, *Draft Federal Information Processing Standards Publication 183*, 21 December 1993. http://www.idef.com/pdf/idef0.pdf
Upward, Frank, "In Search of the Continuum: Ian Maclean's 'Australian Experience' Essays on Recordkeeping," in S. McKemmish and M. Piggot, eds. *The Records Continuum: Ian Maclean and Australian Archives First Fifty Years* (Sydney: Ancora Press in association with Australian Archives, 1994). Online reprint available at http://www.sims.monash.edu.au/research/rcrg/publications/fuptrc.html
Upward, Frank (1996), "Structuring the Records Continuum, Part One: Post-custodial Principles and Properties," *Archives and Manuscripts* 24(2): 268–285. Online reprint available at http://www.sims.monash.edu.au/research/rcrg/publications/recordscontinuum/fupp1.html

Upward, Frank (1997), "Structuring the Records Continuum, Part Two: Structuration Theory and Recordkeeping," *Archives and Manuscripts* 25(1): 10–35. Online reprint available at http://www.sims.monash.edu.au/research/rcrg/publications/recordscontinuum/fupp2.html

von Trier, Lars and Thomas Vinterberg (1995), "The Vow of Chastity." http://www.dogme95.dk/the_vow/vow.html

Warnow-Blewett, Joan, Joel Genuth and Spencer R. Weart, *AIP Study of Multi-Institutional Collaborations: Final Report. Highlights and Project Documentations* (College Park, MD: Center for History of Physics, American Institute of Physics, 2001). http://www.aip.org/history/pubs/collabs/highlights.pdf

Webb, Colin (2003), *Guidelines for the Preservation of Digital Heritage*. Prepared by the National Library of Australia for the Information Society Division, United Nations Educational, Scientific and Cultural Organization, report no. CI-2003/WS/3. http://unesdoc.unesco.org/images/0013/001300/130071e.pdf

Williams, Alena, "Rhizome.org," in *Permanence Through Change: The Variable Media Approach*. Alan Depocas, Jon Ippolito and Caitlin Jones, eds. (New York: Guggenheim Museum Publications, 2003), 39–41. Online reprint available at http://variablemedia.net/pdf/Permanence.pdf

Williams v. Sprint/United Mgmt. Co., 230 F.R.D. 640 (D. Kan. 2005). http://www.ksd.uscourts.gov/opinions/032200JWLDJW-3333.pdf

Wisconsin Department of Administration (2001), *Administrative Rules: Adm 12, Electronic Records Management- Standards and Requirements*. http://www.legis.state.wi.us/rsb/code/adm/adm.html

Zhou, Yuchai (2005), "Profiling and Visualizing Metadata for Geo-referenced Multimedia Information in a Geospatial Portal: A Case Study for the Cybercartography and the New Economy Project" (Master's thesis, Department of Geography and Environmental Studies, Carleton University, 2005).

GLOSSARY

The InterPARES 2 Project Glossary

December 2007

Preface

This glossary is the authoritative list of terms and phrases and their corresponding definitions that were approved by the Terminology Cross-domain Task Force for use throughout the InterPARES 2 research in both working and published documents. The definitions are drawn from many sources, both internal and external to the InterPARES Project. Due to space restrictions, source citations have been omitted in this print version. For information about definition sources, please refer to the online version of the Glossary.[1] For more information about the development of the Glossary, and the Project's other terminological instruments (i.e., Dictionary, Ontologies), see the Terminology Cross-domain Task Force Report.

One final note about the content of this printed version of the glossary. Most of the terms and phrases that are used in to the InterPARES 2 Project's Chain of Preservation and Business-driven Recordkeeping models have been omitted; particularly those terms and phrases whose definitions really only make sense in the specific contexts within which they are used in the models. Definitions for all terms and phrases used in the two models are provided in the appendices to the models (appendices 14 and 15, respectively).

Glossary

acceptance
: *v.,* An agreement, either by express act or by implication from conduct, to the terms of an offer so that a binding contract is formed.

access
: *n.,* The right, opportunity, or means of finding, using or approaching documents and/or information.

access privileges
: *n.,* The authority to access a system to compile, classify, register, retrieve, annotate, read, transfer or destroy records, granted to a person, position or office within an organization or agency. *See also*: access restrictions.

access privileges code
: *n.,* The indication of the person, position or office authorized to annotate a record, delete it, or remove it from the system.

access restrictions
: *n.,* The authority to read a record, granted to a person, position or office within an organization or agency. *See also*: access privileges.

access restrictions code
: *n.,* The indication of the person, position or office authorized to read a record.

access rights
: *See*: access privileges; access restrictions

[1] Available in the InterPARES 2 Terminology Database at http://www.interpares.org/ip2/ip2_terminology_db.cfm.

accessibility
n., The availability and usability of information.

accession
v., To take legal and physical custody of a body of records and to document it in a register.
n., A body of records formally accepted into custody as a unit at a single time.

accession record
n., A record documenting the preserver's acceptance of responsibility for preserving a clearly identified set of records.

accessioned records
n., Acquired records whose authenticity and feasibility of preservation have been confirmed and have been registered in an accession record.

accountability
n., The obligation to answer for actions for which one is responsible.

accrual
n., An acquisition of the records of a creator that is additional to the records of the creator already in the custody of the preserver.

accuracy
n., The degree to which data, information, documents or records are precise, correct, truthful, free of error or distortion, or pertinent to the matter.

acquire
v., To gain possession or control of; to get or obtain.

acquired records
n., Records that are taken into the custody of the preserver from the creator for their permanent preservation.

acquisition
n., An addition to the holdings of an archival repository or records centre.
v., The process of adding to the holdings of an archival repository or records centre.

act
n., The conscious exercise of will by a person aimed to create, maintain, modify or extinguish situations. *Syn.*: action.
n., Legislation that has been made law, especially a statute.

action
See: act

active record
n., A record needed by the creator for the purpose of carrying out the action for which it was created or for frequent reference. *Syn.*: current record. *See also*: inactive record; records lifecycle; semiactive record.

activity
n., A series of acts or actions aimed to one purpose. *See also*: function.

actors
n., Persons who carry out acts or actions.

address
n., A place or location where a particular piece of information is stored, or where an entity can be communicated with.

addressee
n., Person(s) to whom the record is directed or for whom the record is intended.

addressee's name
n., The name of the person to whom the record is directed or for whom the record is intended.

administrative accountability
v., The sort of accountability aimed at non-political and non-legal authorities such as civil servants and top ranking administrators. Involves developing and implementing procedures for carrying out actions and documenting them to ensure that everything is done according to rule and in proper sequence, so that administrators can account at any time precisely for anything that has been done.

administrative context
n., The structure, functions and procedures of the organizational environment in which the creator exists. *See also*: documentary context; juridical-administrative context; procedural context; provenancial context; technological context.

administrative control
n., 1. The exercise of authority over maintenance, use, disposition, and accessibility of current archives to carry on the function for which they were created. *See also*: controlling agency; custodian; custody; legitimate successor; records creator.
n., 2. The means of physically locating the holdings of records centres and archival institutions through numbering and listing. *See also*: intellectual control.

admissibility
n., The quality of being permitted to serve as evidence in trial or hearing or other proceeding.

affix
v., To store on a medium in an unchangeable way. *See also*: save.

affixed
v., Stored on a medium in an unchangeable way.

agency
n., An administrative body having the delegated authority to act competently on behalf of a higher body. Every agency is a juridical person, composed of juridical persons. *See also*: corporate body; organization.

agent
n., One who is authorized to act for or in place of another.

aggregate archival unit
n., An archival unit composed of multiple archival documents, such as a fonds, series or file.

aggregated records
See: archival aggregation; arranged records; records aggregation

alteration
n., An act done to a document after its completion whereby its meaning or language is changed.

American standard code for information interchange
n., (ASCII) The binary code used by most computers to represent in digital format the uppercase and lowercase letters of the Latin script, numerals, and special characters so that they can be sent to, and understood by, other computers and devices such as modems and printers. Each ASCII character consists of seven information bits and one parity bit for error checking. *See also*: unicode; UTF-8.

analogue

a., The representation of an object or physical process through the use of continuously variable electronic signals or mechanical patterns. In contrast to a digitally-encoded representation of an object or physical process, an analogue representation resembles the original.

analogue component

n., An analogue object that is part of one or more analogue documents, requiring a given preservation action. *See also*: digital component.

analogue data

n., The smallest meaningful units of information, expressed as continuous electronic signals or mechanical patterns affixed to an analogue medium. *See also*: digital data.

analogue document

n., An analogue component, or group of components, that is affixed to an analogue medium and is treated and managed as a document. *See also*: digital document.

analogue encoding

v., The use of continuously variable electronic signals or mechanical patterns rather than discrete numeric values (such as those generated by a digital system). *See also*: binary encoding; digital encoding.

analogue medium

n., Physical material, such as a paper, parchment, stone, clay, film or certain types of magnetic audio- and videotape, used for storage of analogue data. *See also*: digital medium.

analogue object

n., A discrete aggregation of one type or class of analogue data (e.g., text, audio, video, image). *See also*: digital object.

analogue record

n., An analogue document that is treated and managed as a record. *See also*: digital record.

analogue system

n., Any system handling analogue data or objects, as opposed to a digital system.

annotation

n., An addition made to a record after it has been created. *See also*: gloss; marginalia; markup [publishing]; notation.

ANSI

Initialism for "American National Standards Institute."

API

Initialism for "application programming interface."

application

n., Computer software that allows the user to process data or perform calculations necessary to achieve a desired result, as opposed to the operating system designed to control the computer's hardware and run all other programs.

application profile

n., A fully conforming instantiation of an element set for a particular community, created to adapt an element set into a package tailored to the functional requirements of a particular application while retaining interoperability with the base standard. Can involve mixing and matching terms from multiple standards to meet the descriptive needs of a particular project or service.

application programming interface

n., (API) A set of routines, protocols and tools for building software applications; specifically, establishing the interface (calling conventions) by which a software application accesses the operating system and other services.

appraisal

n., The process of assessing the value of records for the purpose of determining the length and conditions of their preservation.

appraisal decisions

n., Determinations of the retention periods and disposition of records, including the terms and conditions of transfer from the creator to the preserver.

appraisal strategy

n., The rules and conventions of the entity responsible for permanent preservation that govern the appraisal of records.

appreciation

n., [diplomatics] An intrinsic element of documentary form that comprises a wish for the realization of the content of the document.

architecture

n., [computing] Design, the way components fit together. The term architecture can refer to either hardware or software, or to a combination of hardware and software. The architecture of a system always defines its broad outlines, and may define precise mechanisms as well. *See also*: open architecture.

archival aggregation

n., The whole of the archival documents comprising an aggregate archival unit. *See also*: arranged records.

archival arrangement

n., The process of identifying archival documents as they belong to groupings within a fonds. *See also*: archival unit.

archival bond

n., The network of relationships that each record has with the records belonging in the same records aggregation.

archival date

n., The date on which a record is set aside by inserting it in the records aggregation to which it belongs and classifying it.

archival description

n., The creation of an accurate representation of a unit of description and its component parts, if any, by capturing, analyzing, organizing and recording information that serves to identify, manage, locate and explain archival materials and the context and records systems which produced it. *Syn.*: description. *See also*: administrative control; archival documentation; described records; descriptive instrument; intellectual control.

archival description system

n., The set of descriptive instruments that provide intellectual and physical control over the records of an archival institution or program. Includes, but is not limited to, guides, inventories, indexes, repository locators.

archival document

See: record. *See also*: archives [records].

archival documentation

n., The information provided by a creator and the repository which provides enough information to establish the provenance, history and context of acquired records and to enable its use by others. *See also*: administrative context; archival description; documentary context; juridical-administrative context; procedural context; provenancial context; technological context.

archival fonds

See: fonds

archival framework

n., The whole of the policies, strategies and procedures, based on archival concepts, principles and methods, that control recordkeeping and preservation.

archival methodology

n., The application of principles and procedures articulated and developed from archival theory. The analysis, recording and maintenance of the links between the function of the information recorded on the one hand and its form, structure and provenancial context on the other.

archival preservation

See: records preservation

archival preservation system

See: records preservation system

archival science

n., A systematic body of knowledge that supports the practice of appraising, acquiring, authenticating, preserving, and providing access to recorded materials.

archival unit

n., Any of the groupings of archival documents comprising a fonds, as delineated during the process of archival arrangement. An archival unit can be a fonds, series, file, item or variation thereof, depending on institutional standards. A unit may be divided into sub-units for the purposes of archival description. The smallest (and therefore indivisible) archival unit is the item. *See also*: aggregate archival unit; archival aggregation; arranged records.

archive

v., To save digital data, documents, and records, typically those that are not current, offline.

archives

n., [institution] An agency or institution responsible for the preservation and communication of records selected for permanent preservation.

n., [place] A place where records selected for permanent preservation are kept.

n., [records] The whole of the documents made and received by a juridical or physical person or organization in the conduct of affairs, and preserved. *Syn.*: fonds.

archivist

n., A professional educated in archival science and/or responsible for the administration of archives. *See also*: designated records preserver; records manager; trusted custodian; trusted records officer.

arranged records

n., Records of a creator that have been identified as to their provenance and relationships according to the concepts and principles of archival arrangement.

arrangement

See: archival arrangement.

ASCII
Initialism for "American standard code for information interchange."

assessments of authenticity
n., The determination of whether a document has all the formal elements that is was supposed to present when first made or received and set aside.

attachment
n., A document that, on its being physically connected to a record by an act, becomes part of that record.

attestation
n., [diplomatics] An intrinsic element of documentary form that comprises the written validation of a record (by subscription, superscription or any identifying sign having such function) by those who took part in the issuing of it (author, writer, countersigner) and by witnesses to the action or to the signing of the record.

attributes
n., [diplomatics] The characteristics that uniquely identify a record.

audiovisual
a., Of or relating to the use of sound and/or visual images to present information.

audiovisual record
n., A record in pictorial and/or aural form, regardless of format. *See also*: aural record.

audit trail
n., Documentation of all the interactions with records within an electronic system in which any access to the system is recorded as it occurs.

aural record
n., A record which is a representation of words, music, or any other manifestation of sound that can only be perceived by hearing it. *See also*: audiovisual record; electro-acoustic record.

authentic copy
n., A copy certified by an official authorized to execute such a function, so as to render it legally admissible in court. *See also*: conformed copy; copy in form of original; imitative copy; pseudo-copy; simple copy.

authentic record
n., A record that is what it purports to be and that is free from tampering or corruption. *See also*: authenticated record; authoritative record; complete record; effective record; perfect record; reliable record.

authenticate
v., To declare, either orally, in writing or by affixion of a seal, that an entity is what it purports to be, after having verified its identity. *See also*: certificate of authenticity.

authenticated record
n., A record whose authenticity has been declared at a specific point in time by a juridical person entrusted with the authority to make such a declaration (e.g. public officer, notary, certification authority). *See also*: authentic record.

authentication
n., A declaration of a record's authenticity at a specific point in time by a juridical person entrusted with the authority to make such a declaration (e.g., public officer, notary, certification authority). *See also*: certificate of authenticity.

authenticity

n., The trustworthiness of a record as a record; i.e., the quality of a record that is what it purports to be and that is free from tampering or corruption. *See also*: identity; integrity.

authenticity requirement

n., The specification of the elements of form and context that need to be preserved in order to maintain the authenticity of a given type of electronic record. *See also*: baseline authenticity requirements; benchmark authenticity requirements; presumption of authenticity.

author

n., The physical or juridical person having the authority and capacity to issue the record or in whose name or by whose command the record has been issued.

authoritative copy

n., The instantiation of a record that is considered by the creator to be its official record and is usually subject to procedural controls that are not required for other instantiations.

authoritative record

n., A record that is considered by the creator to be its official record and is usually subject to procedural controls that are not required for other copies. The identification of authoritative records corresponds to the designation of an office of primary responsibility as one of the components of a records retention schedule. *See also*: authentic record; authenticated record; complete record; effective record; perfect record; reliable record.

authoritative version

n., The version of a record that is considered by the creator to be its official record and is usually subject to procedural controls that are not required for other versions.

authority

n., The right or permission to act legally on another's behalf; esp., the power of one person to affect another's legal relations by acts done in accordance with the other's manifestations of assent; the power delegated by a principal to an agent.

back-end database

n., A database that contains and manages data for an information system, distinct from the presentation or interface components of that system.

back up

v., To make a copy of a data file for the purpose of system recovery.

backup

n., A copy of a data file made for the purpose of system recovery.

backward compatibility

n., The ability of a software application or a system to share data or commands with older versions of itself, or sometimes other older applications or systems, particularly applications or systems it intends to supplant. Sometimes backward compatibility is limited to being able to read old data but does not extend to being able to write data in a format that can be read by old versions. *See also*: conversion; preservation strategy; software re-engineering.

baseline authenticity requirements

n., The minimum conditions necessary to enable the preserver to attest to the authenticity of copies of a creator's digital records in the custody of the preserver. *See also*: authenticity requirement; benchmark authenticity requirements.

basic copy

n., A duplicate of a document saved in the file format in which it was originally created or in which it was last used and saved, thus making it more immediately accessible and human-readable in the creator's usual desktop environment.

benchmark authenticity requirements

n., The conditions that serve as a basis for the preserver's assessment of the authenticity of a creator's digital records during appraisal. *See also*: authenticity requirement; baseline authenticity requirements; presumption of authenticity.

best practice

n., In the application of theory to real-life situations, a procedure that, when properly applied, consistently yields superior results and is therefore used as a reference point in evaluating the effectiveness of alternative methods of accomplishing the same task. *See also*: guideline; policy; rule; standard.

binary code

n., A code made up of the digits 0 and 1, called bits, transmitted as a series of electrical pulses (0 bits at low voltage and 1 bits at higher voltage). *See also*: binary encoding.

binary encoding

n., The process of converting data into electronic signals for computer storage and processing purposes. *See also*: analogue encoding; byte-serialized encoding; digital encoding.

bit

n., The smallest unit of data (represented by 0 or 1) that a computer can hold in its memory. *Syn.*: binary bit. *See also*: bitmap; bitstream; byte; byte stream; parity bit.

bitmap

n., A digital representation composed of dots arranged in rows and columns, each represented by a single bit of data that determines the value of a pixel in a monochrome image on a computer screen. In a gray scale or color image, each dot is composed of a set of bits that determine the individual values of a group of pixels that in combination create the visual impression of a specific shade or hue. Also spelled "bit map."

bitstream

n., Digital data encoded in an unstructured sequence of binary bits that are transmitted, stored or received as a unit. Also spelled "bit stream." *See also*: byte stream; data stream.

bounded variability

n., The changes to the form and/or content of a digital record that are limited and controlled by fixed rules, so that the same query, request or interaction always generates the same result. *See also*: fixed form; fixity; stable.

business process

n., A series of rules that governs the carrying out of a transaction.

byte

n., An element in the machine data hierarchy larger than a bit and usually smaller than a word; now nearly always eight bits and the smallest addressable unit of digital storage. A byte typically holds one character.

byte stream

n., a bitstream in which data (binary bits) are grouped into units called bytes. *See also*: data stream.

byte-serialized encoding

v., The process of converting a digital object's bitstream state to a byte stream state.

calling convention
n., In computer application programming, the arrangement of arguments (i.e., values or references) for a procedure or function call.

capture
v., To save a particular instantiation or state of a digital component or group of components.

captured document
n., A made or received digital document that is saved by the creator.

CD
Initialism for "compact disc."

CEN
Initialism for "Comité Européen de Normalisation" (European Committee for Standardization).

central processing unit
n., (CPU) The hardware component of a computer that houses the circuitry for storing and processing data according to instructions contained in the programs installed on it.

certificate of authenticity
n., A declaration by the creator or preserver that one or more reproduced or reproducible digital records is authentic. *See also*: authenticate; authentication.

chain of preservation
n., A system of controls that extends over the entire lifecycle of records in order to ensure their identity and integrity over time. *See also*: administrative control; intellectual control; recordkeeping; records preservation; record-making; unbroken custody.

character
n., One of a set of symbols, such as letters or numbers, that are arranged to express data and/or information. *See also*: byte.

character set
n., A group of symbols used in computing to print and display text electronically. *See also*: ASCII; unicode; UTF-8.

chronological date
n., The date (and, possibly, the time of day) of a record, included in the record by its author, or by the electronic system on the author's behalf, in the course of its compilation. *Syn.*: date of compilation. *See also*: topical date.

ciphertext
n., Text or other data in encrypted form; the product of plaintext after encryption. *See also*: encoding.

classification
n., The systematic organization of records in groups or categories according to methods, procedures, or conventions represented in a plan or scheme.

classification code
n., A series of alphabetical, numerical, or alphanumerical symbols used to identify the record in its documentary context.

classification scheme
n., A plan for the systematic identification and arrangement of business activities and records into categories according to logically structured conventions, methods and procedural rules.

classified record
n., A record, made or received by the creator, which has been given a classification code based on the classification scheme.

classify
v., To arrange systematically in groups or categories according to established criteria.

clause
n., [diplomatics] Any of several textual formulae found after or within the disposition of a document that ensure the execution of the act, avoid its violation, guarantee its validity, preserve the rights of third parties, attest the execution of the required formalities, or indicate the means employed to give the document probative value. *See also*: complimentary clause; final clause.

compact disc
n., (CD) A type of single-sided, optical digital medium that is capable of storing approximately 700 megabytes of digital data on one continuous, microscopic, spiral track or groove that is cut and read by a laser beam. Its logical format specifications are governed by the ISO 9660 and 13490 standards. *See also*: digital videodisc; non-volatile storage; optical disk.

compatibility
n., The ability of different devices or systems (e.g., programs, file formats, protocols, even programming languages) to work together or exchange data without modification. *See also*: cross-platform; interoperability.

competence
n., A sphere of functional responsibility entrusted to a physical or juridical person.

competent person
n., The physical or juridical person given the authority and capacity to act within a specific sphere of responsibility.

compile
v., To gather and put together pieces of existing data or information from various sources in an orderly, structured whole. *See also*: compose.

complete record
n., A record that contains all the elements required by the creator and the juridical system for it to be capable of generating consequences. *See also*: authentic record; complete record; effective record; perfect record; reliable record.

completeness
n., The characteristic of a record that refers to the presence within it of all the elements required by the creator and the juridical system for it to be capable of generating consequences. With primitiveness and effectiveness, a quality presented by an original record.

compliant
v., Ensuring that the requirements of laws, regulations, industry codes and organizational standards are met.

complimentary clause
n., [diplomatics] A brief formula expressing respect, such as “sincerely yours” or “yours truly”. *See also*: clause.

component
n., Uniquely identifiable input, part, piece, assembly or subassembly, system or subsystem, that (1) is required to complete or finish an activity, item, or job, (2) performs a distinctive and necessary function in the operation of a system, or (3) is intended to be included as a part of a finished, packaged, and labeled item. Components are usually removable in one piece and are considered indivisible for a particular purpose or use.

compose
v., To compile, formulate, generate or write the content of a document in an orderly, structured manner.

compression
v., The (re)coding of digital data to save storage space or transmission time. *See also*: lossless compression; lossy compression.

conformed copy
n., An exact copy of a document on which has been written explanations of things that could not or were not copied; e.g., written signature might be replaced on conformed copy with notation that it was signed by the person whose signature appears on the original. *See also*: authentic copy; copy in form of original; imitative copy; pseudo-copy; simple copy.

content
n., The message contained in the body of a record.

context
n., The framework in which a record is created, used, and maintained.

controlling agency
n., The agency exercising administrative control (1) over archival documents. *See also*: custody; records creator.

conversion
n., The process of transforming a digital document or other digital object from one format, or format version, to another one. *See also*: backward compatibility; preservation strategy; software re-engineering; transformative migration; upgrade [*v.*].

conversion of records
n., The process of converting records in the usual and ordinary course of business (otherwise the activity is not conversion but creation) for purposes of security, disaster prevention, conservation, overcoming technology obsolescence, ensuring compatibility with a different hardware or software configuration or generation, or compacting the information, while leaving intact their intellectual form. *See also*: migration of records; refreshing of records; transformative migration of records.

copy
n., The duplicate of an object, resulting from a reproduction process. *Syn.*: duplicate.

copy in form of original
n., A copy identical to the original and having the same effects, but generated subsequently. *See also*: authentic copy; conformed copy; imitative copy; pseudo-copy; simple copy.

copy of outgoing document
n., The reproduction of a document sent to an external juridical or physical person, which is simultaneously set aside as a record by the creator, usually in a recordkeeping system.

corporate body
n., An organization or group of persons created by statute that is identified by a particular name and that acts as a legal entity. *See also*: agency; organization.

correct

a., [diplomatics] The state of a record whose content and form fulfill the requirements set by the creator and the juridical system for the record to reach the consequences or produce the effects for which it was intended. With pertinent, precise and truthful, a component of accuracy.

corroboration

n., [diplomatics] An intrinsic element of documentary form that comprises an explicit statement of the means used to validate a record.

corruption

n., A change in the form, content or attributes of a record such that the record one is accessing can no longer be considered the one that it purports to be.

countersigner

n., The person who, by signing a record, assumes responsibility for the regularity of the procedure generating it and for its forms.

CPU

Initialism for "central processing unit."

created record

n., A made or received document declared a record and set aside for action or reference, usually in a recordkeeping system. *See also*: made record; record creation; record-making.

creation procedure

v., The procedure governing the formation of the record and/or its participation in the act.

creator

See: records creator

cross-platform

The capability of software or hardware to run identically on different platforms; facilitated by the adoption of open-standard products and formats. *Syn.*: platform-independent; platform-neutral. *See also*: compatibility; interoperability.

crosswalk

n., A chart or table that represents the semantic mapping of fields or data elements in one metadata standard to fields or data elements in another standard that has a similar function or meaning.

cryptographic system

n., Any computer system that involves cryptography.

cryptography

n., The practice and study of protecting information by transforming it (encrypting it) into an unreadable format, called cipher text. Only those who possess a secret (private) key can decipher (or decrypt) the message into plain text.

cryptosystem

Shorthand for "cryptographic system."

current record

See: active record

custodian

n., A person or institution that has charge or custody (of a child, property, papers, or other valuables).

custody

n., The basic responsibility for guardianship of records/archives based upon their physical possession but not necessarily implying legal title. *See also*: administrative control; controlling agency.

DAT

Initialism for "digital audio tape."

data

n., The smallest meaningful units of information.

data element

n., A discrete component of data. *See also*: field.

data format

n., The organization of data within files, usually designed to facilitate the storage, retrieval, processing, presentation, or transmission of the data by software.

data grid

n., The registration of digital entities into a logical name space. Manipulations of registered material can then be automated through any standard computer application programming interface (API).

data hierarchy

n., The system of data objects which provide the methods for information storage and retrieval. Broadly, a data hierarchy may be considered to be either natural or machine.

data model

n., The product of the database design process that aims to identify and organize the required data logically and physically. A data model says what information is to be contained in a database, how the information will be used, and how the items in the database will be related to each other.

data processing

v., The systematic performance of a single operation or sequence of operations by one or more central processing units on data converted to machine-readable format to achieve the result for which the computer program that controls the processing was written.

data restoration

n., The process of recovering data or records as bits from a failed, damaged, degraded or obsolete digital medium, followed by steps to restore the intelligibility of the recovered data or records. *Syn.*: digital archaeology. *See also*: preservation strategy.

data stream

n., A sequence of digitally encoded signals used to represent information in transmission. Also spelled "datastream." *See also*: bitstream; byte stream.

data type

n., The representation of information according to preset specifications (e.g., plain text files, HTML, TIFF, etc.).

database

n., A collection of data items and links between them, structured in a way that allows it to be accessed, manipulated and extracted by a number of different applications programs or query languages.

date of compilation

See: chronological date.

date of document

n., The particular place (topical date) and/or time (chronological date) of the compilation and/or issuing of recorded information (the document).

date of receipt

n., The date the record is received by the agency to which it was sent.

date of record

n., The date assigned to it by the author.

date of transmission

n., The date the record leaves the space in which it was generated.

de facto standard

n., A standard not issued by any official standards-setting body, but nevertheless widely used and recognized by its users as a standard.

de jure standard

n., A standard issued by an official standards-setting body, whether national (e.g., ANSI), multi-national (e.g., CEN) or international (e.g., ISO).

declared record

n., An identified document, made or received by the creator, which has been classified and registered.

decompression

v., Returning a compressed image or compressed data to its uncompressed form. Some compression methods lose information so that the uncompressed image or data is not equivalent to the original.

degree of perfection

n., The completeness, primitiveness and effectiveness (enforceability) of a record. The three degrees of perfection for a record are draft, original and copy. *See also*: status of transmission.

described records

n., Arranged records for which information about their nature, make-up and contexts (juridical-administrative, provenancial, procedural, documentary and technological) are recorded to facilitate administrative and intellectual control. *See also*: archival description; descriptive instrument.

description

See: archival description

description of action or matter

n., Presentation of the ideal motivation (preamble) and the concrete reason (exposition) for the action as well as the action or matter itself (disposition) as expressed in the content of the record.

descriptive instrument

n., A tool prepared in the course of archival description and indexing of records for the purposes of administrative and intellectual control. *See also*: described records.

designated records preserver

n., The entity responsible for taking physical and legal custody of and preserving (i.e., protecting and ensuring continuous access to) authentic copies of a creator's inactive records. The role of the designated records preserver should be that of a trusted custodian for a creator's records. *Syn.*: designated preserver; preserver. *See also*: archivist; legitimate successor; records manager; trusted records officer.

digital

a., The representation of a physical process through discrete, binary values. In contrast to an analogue representation of an object or physical process, a digitally-encoded representation does not resemble the original.

digital archaeology

See: data restoration

digital audio tape

n., (DAT) A type of magnetic digital medium that can store up to 4 gigabytes of digital data per cassette by using helical scan recording. *See also*: digital linear tape; non-volatile storage.

digital component

n., A digital object that is part of one or more digital documents, and the metadata necessary to order, structure or manifest its content and form, requiring a given preservation action. *See also*: analogue component.

digital data

n., The smallest meaningful units of information, expressed as binary bits that are digitally encoded and affixed to a digital medium. *See also*: analogue data.

digital document

n., A digital component, or group of digital components, that is saved and is treated and managed as a document. *See also*: analogue document.

digital encoding

v., The use of discrete numeric values (such as the binary values 0 and 1) rather than a continuous spectrum of values (such as those generated by an analogue system). *See also*: analogue encoding; binary encoding.

digital entity

n., A real or abstract digital construct. *See also*: digital object.

digital format

n., The byte-serialized encoding of a digital object that defines the syntactic and semantic rules for the mapping from an information model to a byte stream and the inverse mapping from that byte stream back to the original information model. In most contexts, digital format is used interchangeably with digital file-related concepts such as file format, file wrapper, file encoding, etc. However, there are some contexts, "such as the network transport of formatted content streams or consideration of content streams at a level of granularity finer than that of an entire file, where specific reference to "file" is inappropriate." *Syn.*: digital presentation. *See also*: file format.

digital linear tape

n., (DLT) A type of magnetic digital medium that can store up to 35 gigabytes of digital data per cassette by using longitudinal recording. *See also*: digital audio tape; non-volatile storage.

digital medium

n., Physical material, such as a CD, DVD, DAT or hard disk, used for storage of digital data. *See also*: analogue medium.

digital object

n., A discrete aggregation of one or more bitstreams and the metadata about the properties of the object and, if applicable, methods of performing operations on the object. *See also*: analogue object.

digital presentation

See: digital format

digital preservation

v., The specific process of maintaining digital materials during and across different generations of technology over time, irrespective where they reside. *See also*: persistent object preservation; records preservation.

digital record

n., A digital document that is treated and managed as a record. *See also*: analogue record.

digital signature

n., An electronic signature based on public key cryptography.

digital system

n., Any system handling binary data, as opposed to an analogue system.

digital timestamp

n., A binary code attached to a record indicating the time that an event occurred, such as creation, receipt, reading, modification or deletion.

digital videodisc

n., (DVD) A type of single- or double-sided, optical digital medium that is capable of storing from 4.7 to 8.5 gigabytes of digital data on two continuous, microscopic, spiral tracks or grooves that are cut and read by a laser beam. Its logical format specifications are governed by the Universal Disk Format (UDF) specification. *Syn.*: digital versatile disc. *See also*: compact disc; non-volatile storage; optical disk.

diplomatics

n., The discipline which studies the genesis, forms and transmission of archival documents, and their relationship with the facts represented in them and with their creator, in order to identify, evaluate, and communicate their true nature.

directive

n., An order or instruction, especially one issued by an agency, corporate body, organization or other central authority. *See also*: best practice; guideline; policy; rule; standard.

disposition

n., [archives] Records' final destruction or transfer to an archives as determined by their appraisal.

n., [diplomatics] An intrinsic element of documentary form that comprises the core of the text of a document narrating the expression of the will of the author and the action of the record.

dispositive

a., (of a document) Constituting a juridical act. *See also*: narrative; probative; supporting.

dispositive record

n., A retrospective record whose purpose is to put into existence an act, the effects of which are determined by the writing itself; that is, the written form of the record is the essence and substance of the act. With enabling, instructive, narrative, probative and supporting, one of six functional categories of records.

DLT

Initialism for "digital linear tape."

document

n., An indivisible unit of information constituted by a message affixed to a medium (recorded) in a stable syntactic manner. A document has fixed form and stable content. *See also*: analogue document; digital document; record; work; written record. *See also*: analogue document; archival document; attachment; captured document; copy of outgoing document; digital document; HTML document; identified document; incoming document; internal document; made document; manifested digital document; outgoing document; potential record; received document; record; schema document; sent document; SGML document; SGML-compliant document; stored digital document; work; written record; XML document.

document schema

See: schema document

document type definition

n., (DTD) A formal specification that contains or points to the syntactic rules according to which an SGML-compliant document can be composed. *See also*: schema document.

documentary context

n., The archival fonds to which a record belongs, and its internal structure. *See also*: administrative context; juridical-administrative context; procedural context; provenancial context; technological context.

documentary form

n., The rules of representation according to which the content of a record, its administrative and documentary context, and its authority are communicated. Documentary form possesses both extrinsic and intrinsic elements. *Syn.*: documentary presentation. *See also*: form; overall presentation.

documentary presentation

See: documentary form

documentary procedure

n., The body of rules governing the making of an archival document. The more standardized and rigorous the procedure, the more reliable the record is presumed to be.

documentation

n., All material that serves primarily to describe a system and make it more readily understandable, rather than to contribute in some way to the actual operation of the system. Documentation is frequently classified according to purpose; thus for a given system there may be requirements documents, design documents, and so on. In contrast to documentation oriented toward development and maintenance of the system, user documentation describes those aspects of the system that are of interest to end-users.

dossier

n., The aggregation of all the records that participate in the same affair or relate to the same event, person, place, project, or other subject. *Syn.*: file. *See also*: fonds; item; series.

draft

n., A record made for purposes of correction.

DRAM

Initialism for “dynamic random-access memory.”

DTD

Initialism for “document type definition.”

duplicate

See: copy

durable encoding

n., The encoding of digital documents or other digital objects to conform to well-known data processing standards down to the level of encoding bits as ASCII or Unicode UTF-8, and objects as XML. *See also*: normalization; preservation strategy.

DVD

Initialism for "digital videodisc" or "digital versatile disc."

dynamic computing

n., Flexible and adaptable approaches to tailoring computing resources to demands.

dynamic random-access memory

n., (DRAM) A type of semiconductor memory in which the information is stored in capacitors on a metal oxide semiconductor integrated circuit. Due to leakage the capacitor discharges gradually and the memory cell loses the information. Therefore, to preserve the information, the memory has to be refreshed periodically.

dynamic record

n., A record the content of which is dependent upon data that might have variable instantiations and be held in databases and spreadsheets internal or external to the system in which the record is generated. *See also*: experiential record; interactive record.

dynamic system

n., A system linked to particular algorithmic programming and mathematical system capabilities, as expressed in this statement: "The identification of dynamic systems concerns the definition of a mathematical model which behaves like a process solely on the basis of its measurements."

effective record

n., A record capable of reaching the consequences or producing the effects for which it was intended. *See also*: authentic record; authoritative record; complete record; perfect record; reliable record.

effectiveness

a., The characteristic of a record that refers to the presence within it of all the elements required by the creator and the juridical system for it to be capable of reaching the consequences or producing the effects for which it was intended. With completeness and primitiveness, a quality presented by an original record.

e-government

n., The use of information technologies, especially the Internet to improve government services for and interactions with citizens (G2C) , businesses and industry (G2B), and different division of government (G2G) by simplifying processes, and by integrating and eliminating redundant systems.

electroacoustics

n., A science that deals with transformation of sound energy into electric energy and vice versa.

electroacoustic record

n., An aural record generated by an electronic device.

electronic

n., Device or technology associated with or employing low voltage current and solid state integrated circuits or components, usually for transmission and/or processing of analogue or digital data.

electronic address
n., A location of data, usually in main memory or on a disk. You can think of computer memory as an array of storage boxes, each of which is one byte in length. Each box has an address (a unique number) assigned to it. By specifying a memory address, programmers can access a particular byte of data. Disks are divided into tracks and sectors, each of which has a unique address.

electronic record
n., An analogue or digital record that is carried by an electrical conductor and requires the use of electronic equipment to be intelligible by a person.

electronic seal
n., A means of authenticating a record and its author or a means of protecting the confidentiality of the record by ensuring that the record is only opened by the intended addressee. It is a distinct type of electronic signature.

electronic signature
n., A digital mark that has the function of a signature in, is attached to, or is logically associated with a record, and is used by a signatory to take responsibility for, or to give consent to, the content of the record.

electronic system
n., [Electronic Records System] Any information system that produces, processes or stores records by using a computer. Often called an automated information system.

element
n., A fundamental, essential, or irreducible constituent of a composite entity. *See also*: data element; element of form; element set; extrinsic element; intrinsic element; metadata element; record element.

element of form
n., A constituent part of the record's documentary form, visible on the face of the record. It may be either extrinsic, like a seal, or intrinsic, like a subscription.

element set
n., Collections of elements that have at least one characteristic in common.

e-mail
n., An abbreviation of electronic mail, an Internet protocol that allows computer users to exchange messages and data files in real time with other users, locally and across networks.

e-mail attachment
n., A file that is linked to and is transmitted along with an e-mail message. The attached file can be of any type.

emulation
n., The reproduction of the behaviour and results of obsolete software or systems through the development of new hardware and/or software to allow execution of the old software or systems on future computers. *Syn.*: preservation emulation. *See also*: encapsulation; preservation strategy; wrapper.

enabling record
n., A prospective record encoded in machine language that is actively involved in carrying out an action or process. With dispositive, instructive, narrative, probative and supporting, one of six functional categories of records.

encapsulation

n., The process of binding together a digital document or other digital object and the means of providing access to it, normally in a wrapper that describes what it is in a way that can be understood by a wide range of technologies (such as an XML document). *See also*: emulation; preservation strategy.

encoding

n., The representation of symbols in some alphabet by symbols or strings of symbols in some other alphabet. *See also*: analogue encoding; binary encoding; byte-serialized encoding; ciphertext; digital encoding; durable encoding; encryption; metadata encoding scheme.

encryption

n., The conversion of data into a secret code (or of plaintext into ciphertext) for transmission over a public network. *See also*: encoding.

entitling

n., [diplomatics] An intrinsic element of documentary form that comprises the name, title, capacity and address of the physical or juridical person issuing the document, or of which the author of the document is an agent.

entity

n., A real or abstract thing. *See also*: object.

entity type

n., The kind of entity identified or described in or addressed by a metadata schema; for example, fonds, records, agents, recordkeeping business.

eschatocol

n., [documentary form] The final part of the document, which contains the documentary context of the action and the final formulae.

evidence

n., All the means by which any alleged matter of fact, the truth of which is submitted to investigation, is established or disproved.

executed record

n., A record that has participated in the execution phase of an administrative procedure and to which metadata that convey the actions taken during the course of the procedure have been attached, such as priority of transmission, transmission date, time and/or place, actions taken, etc.

execution annotation

n., An addition made to a record, after its creation, as part of the document execution phase of an administrative procedure.

execution phase

n., An administrative procedure constituted by all the actions (validation, communication, notification, publication) that give formal character to the transaction and the resulting record.

experiential record

n., A record produced, used and maintained in an experiential system. *See also*: dynamic record; interactive record.

experiential system

n., A system which immerses the user in a sensory experience.

exposition

n., [diplomatics] An intrinsic element of documentary form that comprises the part of the text of a document that narrates the concrete and immediate circumstances generating the act and/or the document.

extensible markup language

n., (XML) A general-purpose specification for creating custom, cross-platform, text-based, markup languages used both to encode documents and to serialize data; a subset of General Standardized Markup Language (SGML) with use and design similar to Hypertext Markup Language (HTML) but employing user-definable markup tags that indicate the logical structure in addition to the display specifications of data elements. *See also*: document type definition; schema document.

external user

n., Secondary users (e.g., journalists, researchers and other external users) who are allowed to access the records creator's recordkeeping system on-line.

extrinsic element

n., [diplomatics] An element of the documentary form of a record that constitutes its external appearance. The types of extrinsic elements include presentation features, electronic signatures, electronic seals, digital timestamps issued by a trusted third party, and special signs. *See also*: intrinsic element.

fact

n., A condition or event that exists, as distinguished from its effects, consequences or interpretations, such as a state of things or a motion.

feasibility information

n., Assessment of the cost and technical capability required for the permanent preservation of a given body of records.

feasibility of preservation

a., The determination that the digital components conferring identity and ensuring the integrity of the things you want to preserve can indeed be preserved given current and future anticipated preservation capabilities.

field

n., An item of data consisting of a number of characters, bytes, words, or codes that are treated together, e.g., to form a number, a name, or an address. *See also*: data element.

file

v., To set aside a made or received document among the records that participate in the same action/affair or relate to the same person or subject, so that they may be retrieved for action or reference.

n., The aggregation of all the records that participate in the same affair or relate to the same event, person, place, project, or other subject. *Syn.*: dossier. *See also*: fonds; item; series.

file format

n., The organization of data within digital objects, usually designed to facilitate the storage, retrieval, processing, presentation and/or transmission of the data by software. *See also*: digital format.

final clause

n., A formula at the end of the text of a document intended to ensure the execution of the act embodied in the document, or avoid its violation, guarantee its validity, preserve the rights of third parties, attest the execution of required formalities, or indicate the means employed to give the document probative value. *See also*: clause.

first manifestation of a record

n., The documentary form that a record has when it is open for the first time upon receipt or after having been captured and declared as a record. *See also*: original record.

fixed form

n., The quality of a record that ensures its content remains complete and unaltered. *See also*: bounded variability. *See also*: bounded variability; fixity; stable.

fixity

a., The quality of a record that makes it immutable and requires changes to be made by appending an update or creating a new version. *See also*: bounded variability; fixed form; stable.

folder

n., A cover in which non-electronic records, belonging in the same dossier, are loosely kept, usually in chronological order. A dossier may be distributed across a number of folders.

fonds

n., The whole of the documents that a physical or juridical person accumulates by reason of its function or activity; the highest-level archival aggregation. *Syn.*: archives [records]; archival fonds. *See also*: dossier; file; item; series.

form

n., Rules of representation that determine the appearance of an entity and convey its meaning. *See also*: documentary form; format; overall presentation.

format

n., The structure or layout of an entity. *See also*: documentary form; file format; form; overall presentation; wrapper format.

formula perpetuitatis

n., [diplomatics] An intrinsic element of documentary form that comprises a sentence declaring that the rights put into existence by the document are not circumscribed by time.

framework

n., A basic conceptual structure of rules, policies, procedures, tools and mechanisms intended to serve as a support or guide for the design, implementation and maintenance of an integrated system.

freeze

v., To lock an evolving software distribution or document against changes so it can be released with some hope of stability.

function

n., All of the activities aimed to accomplish one purpose, considered abstractly.

functionality

n., The capabilities or behaviours of a computer program, part of a program, or system, seen as the sum of its features.

genuine record

n., A record having actual character, which is not counterfeited, imitated or adulterated, and a definite origin from a certain known source.

genuineness

n., The quality of a record that is truly what it purports to be. The two concepts that allow for the indirect assessment and verification of the genuineness of a record are those of reliability and authenticity.

geographic information system

n., (GIS) A computer-based system consisting of hardware, software, geographic information, procedures and personnel designed to facilitate the efficient capture, storage, maintenance, manipulation, analysis, querying and display of spatially-referenced (geospatial) data from a wide variety of data sources in a wide variety of data formats—such as maps, graphs, photographs, remotely sensed data, tabular data and text—providing an automated link between the geospatial (locational) data and the descriptive (attribute) data, usually in relation to a system of coordinates (latitude, longitude, elevation or depth, etc.).

GIS

Initialism for "geographic information system."

gloss

n., A note explaining a difficult or obscure word or phrase, especially such a note made between lines or in the margin of a document. *See also*: annotation; markup [publishing]; marginalia; notation.

graphic record

n., A record that represents an object or outline of a figure, plan, or sketch by means of lines. A representation of an object formed by drawing.

guideline

n., A statement or other indication of policy or procedure by which to determine a course of action, accomplish a given task or achieve a set of goals and objectives, formulated by a body with authority to speak on the subject but less binding than a directive or formal standard. *See also*: best practice; rule.

handling office

n., The office (or officer) formally competent for carrying out the action to which the record relates or for the matter to which the record pertains. *Syn.*: handling person.

handling person

See: handling office

hard disk

n., A magnetic digital medium, composed of one or more flat, circular plates (platters) of a hard material capable of storing a large quantity of digital data, which (typically) resides permanently within a computer. *Syn.*: hard drive. *See also*: non-volatile storage.

hearsay rule

n., A legal provision excluding testimony that is based on second-hand, rather than personal, knowledge (hearsay).

hierarchy

n., A group of related entities, or information about the entities, within a system arranged in a graded order, typically from the most general to the most specific.

historical accountability

n., A need to provide and receive explanation and understanding from one generation to another. Groups of individuals often derive their cohesiveness, legitimacy and the authority for their actions from their understanding and evaluation of the past.

HTML

Initialism for "Hypertext Markup Language."

HTML document

n., An SGML-compliant digital document encoded using Hypertext Markup Language (HTML) in conformance with the syntactic rules described in a Document Type Definition (DTD) or a schema document. *See also*: SGML document; XML document.

human-readable format

n., A document or code that can be read by a human being, with or without the aid of magnification, as opposed to one in a format that can be read only by a computer. *See also*: machine-readable format; machine language.

hypertext

n., A method of presenting digital information that allows related files and elements of data to be interlinked, rather than viewed in linear sequence; usually differentiated from the normal text in a document by a different colour, by underlining, or by both.

hypertext markup language

n., (HTML) A cross-platform, text-based, markup language used to define a single, fixed type of document with markup tags that structure the layout, styling and display of text and provide some provision for hypertext and multimedia; a subset of General Standardized Markup Language (SGML). *See also*: extensible markup language.

IDEF0

Acronym for "integrated definition function modeling method."

identified document

n., A made or received document to which the identity metadata (e.g., persons, actions and dates of compilation) have been attached.

identifier

n., Terms, such as acronyms, projects, proper names of persons, geographical locations, the number of a patent's specification or of a national standard, or any part or a bibliographical description, test names, and trade names which provide subject indexing, in addition to descriptors.

identity

n., The whole of the characteristics of a document or a record that uniquely identify it and distinguish it from any other document or record. With integrity, a component of authenticity.

imitative copy

n., A copy that reproduces both the content and form of the record, but in such a way that it is always possible to tell the copy from the original. *See also*: authentic copy; conformed copy; copy in form of original; pseudo-copy; simple copy.

inactive record

n., A record that is no longer used in the day-to-day course of business, but which may be kept and occasionally used for legal, historical, or operational purposes. *Syn.*: non-current record. *See also*: active record; obsolete record; records lifecycle; semiactive record.

incoming document

See: received document

incoming record

See: received record

indication of action

n., The subject line(s) and/or the title at the top of a record.

indication of attachments

n., The mention of autonomous items that have been linked inextricably to the record before transmission (i.e., added during its execution) for it to accomplish its purpose.

information
n., An assemblage of data intended for communication either through space or across time.

information system
n., A computer-based system with the defining characteristic that it provides information to users in one or more organizations. Information systems are thus distinguished from, for example, real-time control systems, message-switching systems, software engineering environments, or personal computing systems.

initialization
n., A group of commands that initialize a device, such as a printer.

inscription
n., [diplomatics] An intrinsic element of documentary form that comprises the name, title and address of the addressee of the document and/or the action.

instantiation
n., To represent by an instance.

instructive record
n., A prospective record that contains instructions about executing an action or process. With dispositive, enabling, narrative, probative and supporting, one of six functional categories of records.

intact record
n., A record that has had no relevant part removed or destroyed.

integrated business and documentary procedures
n., Procedures for carrying out the creator's business that have been linked to a scheme or plan for organization of the creator's records.

integrated definition function modeling method
n., (IDEF0) A method designed to model the decisions, actions, and activities of an organization or system.

integrity
n., The quality of being complete and unaltered in all essential respects. With identity, a component of authenticity.

intellectual content
See: intellectual form

intellectual control
n., The control established over archival material by documenting in finding aids its provenance, arrangement, composition, scope, informational content and internal and external relationships. *See also*: administrative control.

intellectual form
n., [diplomatics] The whole of the formal attributes of the record that represent and communicate the elements of the action in which the record is involved and of its immediate context, both documentary and administrative. *See also*: physical form.

intellectual property rights
n., (IPR) The rights of individuals or organizations to control the use or dissemination of ideas or information. They include copyright, trademarks and patents.

interactive record
n., A record with variable content or form that is dependent on user input that is often based on earlier content. *See also*: dynamic record; experiential record.

interactive system
n., A system in which each user entry causes a response from or an action by the system, by virtue of automated reasoning based on data from its apparatus.

interactivity
v., An expression of the extent that in a given series of communication exchanges, any third (or later) transmission (or message) is related to the degree to which previous exchanges referred to even earlier transmissions.

internal document
n., A document that is to transmitted across space or through time only within the creating organization, rather than to an external party.

internal record
n., A record, such as a memorandum or report, which is transmitted across space or through time only within the creating organization, rather than to an external party. *See also*: created record; made record; outgoing document; received record.

interoperability
n., The ability of one application/system to communicate or work with another. *See also*: compatibility; cross-platform.

intrinsic element
n., [diplomatics] An element of the documentary form of a record that constitutes its internal composition and that conveys the action in which the record participates and its immediate context. *See also*: extrinsic element.

inventory
n., A descriptive instrument that represents the records of a fonds in their hierarchical structure and arrangement, and illustrates the administrative history of their creator, their custodial history and their administrative and documentary context.

invocation
n., [diplomatics] An intrinsic element of documentary form that comprises the mention of God in documents issued by religious bodies.

IPR
Initialism for "intellectual property rights."

ISO
Initialism for "International Organization for Standardization."

item
n., The smallest indivisible archival unit, which represents the smallest intellectual entity within a fonds no longer usefully subdivisible for descriptive purposes. *See also*: dossier; file; fonds; series.

juridical-administrative context
n., The legal and organizational system in which the creating body belongs. *See also*: administrative context; documentary context; procedural context; provenancial context; technological context.

juridical person
n., An entity having the capacity or the potential to act legally and constituted either by a succession or collection of physical persons or a collection of properties.

juridical system
n., A social group that is organized on the basis of a system of rules and that includes three components: the social group, the organizational principle of the social group, and the system of binding rules recognized by the social group.

LAN
Initialism for “local area network.”
language
n., [diplomatics] An extrinsic element of documentary form that comprises the expression and organization of ideas and discourse including composition, style, rhythm, and vocabulary.
legitimate successor
n., Either the person or organization acquiring the function(s) from which the records in question result and the records themselves, or a designated records preserver. *See also*: administrative control; controlling agency.
lifecycle
See: records lifecycle
list of annotations
n., Recorded information about additions made to a record after it has been created.
list of format changes
n., Recorded information about modifications to a record’s documentary form or digital format after it has been created.
literary warrant
n., The mandate from law, professional best practices, professional literature, and other social sources requiring the creation and continued maintenance of archival description and other metadata supporting the accuracy, reliability, authenticity and preservation of records.
local area network
n., (LAN) A data communications network that is geographically limited (typically to a one kilometre radius) allowing easy interconnection of terminals, microprocessors and computers within adjacent buildings. *See also*: wide area network (WAN) and metropolitan area network (MAN).
logical format
n., The organized arrangement of data on a digital medium that ensures file and data control structures are recognizable and recoverable by the host computer operating system. Two common logical formats for files and directories are ISO 9660/13490 for CDs, and Universal Disk Format (UDF) for DVDs.
logical namespace
n., A set of names in which all names are unique.
logical structure
n., The syntactic organization of data elements in an SGML-compliant document. *See also*: physical structure.
longevity
n., Long life; long duration of existence.
lossless compression
n., A term describing a data compression algorithm which retains all the information in the data, allowing it to be recovered perfectly by decompression. *Opp.*: lossy compression. *See also*: compression.
lossy compression
n., A term describing a data compression algorithm which actually reduces the amount of information in the data, rather than just the number of bits used to represent that information. *Opp.*: lossless compression. *See also*: compression.

machine code
See: machine language

machine hierarchy
n., A data hierarchy that reflects the facilities of the computer, both hardware and software. *See also*: natural hierarchy.

machine language
n., A set of instructions for a specific central processing unit, designed to be usable by a computer without being translated. *Syn.*: machine code. *See also*: machine-readable format.

machine-readable format
n., Data in a form that can be recognized, accepted, and interpreted by a machine, such as a computer or other data processing device, whether created in such a form or converted from a format that a machine cannot read. *See also*: human-readable format; machine language.

made document
n., A document composed or compiled by the creator. *See also*: received document.

made record
n., A made document declared a record and set aside for action or reference, usually in a recordkeeping system. *See also*: created record; internal record; received record; record creation; record-making.

maintenance strategy
n., A coherent set of objectives and methods for protecting and maintaining accessibility of authentic copies of digital records through their early stages in the chain of preservation. *See also*: preservation strategy.

MAN
Initialism for "metropolitan area network."

manifest
v., To render a stored digital object in a form suitable for presentation either to a person (i.e., in human-readable form) or to a computer system (i.e., in machine language). *See also*: reconstitute.

manifested digital component
n., A digital component that is visualized or rendered from a stored digital component in a form suitable for presentation within a document either to a person (i.e., in human-readable form) or to a computer system (i.e., in machine language).

manifested digital document
n., A digital document that is visualized or rendered from a stored digital document and/or stored digital component(s) in a form suitable for presentation either to a person (i.e., in human readable form) or to a computer system (i.e., in machine language).

manifested digital record
n., A manifested digital document that is treated as a record. *Syn.*: presented digital record. *See also*: reproduced digital record.

marginalia
n., Notes made in the margin of a document. *Syn.*: marginal notes. *See also*: annotation; gloss; markup [publishing]; notation.

mark-up
v., To add codes (markup tags) to a digital document to give semantic structure to the content. *Syn.*: tag. *See also*: document schema; document type definition; markup language.

markup language

n., A computer-processable encoding language and associated rules that can be used to mark-up or tag SGML-compliant documents to indicate their logical structure, layout, display and styling. *See also*: extensible markup language; hypertext markup language; standard generalized markup language.

markup tag

n., The markup characters that indicate the start or end of a data element in an SGML-compliant document. A tag serves as an instruction to a processing or reading program, specifying how the data element is defined or displayed. *See also*: document schema; document type definition.

medium

n., The physical material or substance upon which information can be or is recorded or stored. *See also*: analogue medium; digital medium.

n., [diplomatics] An extrinsic element of documentary form that comprises the material carrying the message of a document.

metadata

n., [data] Information that characterizes another information resource, especially for purposes of documenting, describing, preserving or managing that resource.

n., [file] Any file or database that holds information about a document, record, records aggregation or another database's structure, attributes, processing or changes.

metadata element

n., A discrete component of metadata.

metadata element set

n., A grouping of metadata elements along with their attributes, such as name, identifier, definition or relationship to other concepts, collated for a specific purpose, community or domain.

metadata encoding scheme

n., A controlled vocabulary for metadata element values and encoding structures.

metadata instance

n., A particular collection of metadata elements associated with a set of values for those elements.

metadata mapping

n., A formal identification of equivalent or nearly equivalent metadata elements or groups of metadata elements within different metadata schemas, carried out in order to facilitate semantic interoperability.

metadata registry

n., An authoritative source of repository for names, semantics and syntaxes for one or more schemas. An application that uses metadata languages in a form processable by machines to make those languages available for use by both humans and machines.

metadata schema

n., A framework that specifies and describes a standard set of metadata elements and their interrelationships that need to be recorded to ensure the identification of records and their authenticity. Schemas provide a formal syntax (or structure) and semantics (or definitions) for the metadata elements.

metadata schema registry

n., An authoritative resource that can be used to catalogue, describe, document and analyze metadata schemas and versions thereof.

metadata set
See: metadata element set

metropolitan area network
n., (MAN) A data network intended to serve an area the size of a large city. *See also*: local area network (LAN) and wide area network (WAN).

methodology
n., The system of broad principles or rules from which specific methods or procedures may be derived to understand different situations (or solve different problems) within the scope of a particular discipline.

migration
n., The process of moving or transferring digital objects from one system to another. *See also*: refreshing; transformative migration.

migration of records
n., The process of moving records from one system to another to ensure their continued accessibility as the system becomes obsolete, while leaving intact their physical and intellectual forms. *See also*: conversion of records; refreshing; transformative migration of records.

mode of transmission
n., The method of transmission of a record (e.g., by fax).

model
n., A schematic description of a system, theory, or phenomenon that accounts for its known or inferred properties and may be used for further study of its characteristics.

name of action or matter
n., The subject line(s) and/or the title at the top of the record.

name of place of origin
n., The name of the geographic place where the record was generated; included in the content of the record by the author or by the electronic system on the author's behalf.

namespace
n., A collection of names, identified by a URL reference, used as element types and attribute names.

narrative
a., (of a document) Constituting evidence of a juridically irrelevant activity. *See also*: dispositive; probative; supporting.

narrative record
n., A retrospective record constituting written evidence of activities that are juridically irrelevant. With dispositive, enabling, instructive, probative and supporting, one of six functional categories of records.

natural hierarchy
n., A data hierarchy that arises from the alphabet or syntax of the language in which the information is expressed. *See also*: machine hierarchy.

natural person
See: physical person

network
n., A hardware and software data communication system. Networks are often also classified according to their geographical extent: local area network (LAN), metropolitan area network (MAN), wide area network (WAN) and also according to the protocols used.

non-current record
See: inactive record

nonproprietary
a., In reference to hardware technology, software applications and/or file formats, the state of not being protected by trademark, patent or copyright nor owned or controlled solely by one company or institution. *Opp*: proprietary. *See also*: open architecture; open source; open-standard products and formats.

nonrepudiation
n., The capacity of a digital security service of ensuring that a transferred message has been sent and received by the parties claiming to have sent and received the message, and of providing proof of the integrity and origin of data, both in an unforgeable relationship, which can be verified by any third party at any time.

non-volatile storage
n., Computer memory that does not require electrical power to maintain its stored content. *Syn.*: persistent storage; secondary storage. *Opp.*: volatile storage. *See also*: compact disc; digital audio tape; digital linear tape; digital videodisc; hard disk.

normalization
n., The process or creating and/or storing digital documents or other digital objects in a limited number of, often standardized, data or file formats. *See also*: durable encoding; preservation strategy.

notation
n., An annotation or comment in a document, often handwritten marginalia or a gloss. *See also*: markup [publishing].

notification
n., [diplomatics] An intrinsic element of documentary form that comprises the expression of the purport of the document intended to communicate to all who have an interest in the act consigned to the document.

object
n., A tangible entity. *See also*: analogue object; digital object.

obsolescence
n., The process of becoming obsolete or the condition of becoming nearly obsolete.

obsolete record
n., A record no longer in use or no longer useable or useful to the creator for action or reference. *See also*: inactive record.

obsolete technology
n., An out-of-date technological invention no longer in use.

office of primary responsibility
n., The office given the formal competence for maintaining the authoritative version or copy of records belonging to a given class within a classification scheme.

official record
n., A complete, final, and authorized version or instantiation of a record. *See also*: authoritative record.

open architecture
n., [computing] An architecture whose specifications are public. This includes officially approved standards as well as privately designed architectures whose specifications are made public by the designers. *See also*: nonproprietary; open-source; open-standard products and formats.

open-source

n., A computer program in which the source code is available to the general public for use and/or modification from its original design free of charge (open).

a., A method and philosophy for software licensing and distribution designed to encourage use and improvement of software written by volunteers by ensuring that anyone can copy the source code and modify it freely. *Opp*: proprietary. *See also*: nonproprietary; open architecture; open-standard products and formats.

open-standard products and formats

n., Freely available structures, procedures or tools for the uniform creation and description of data. Usually defined and perhaps maintained by a central body, but, unlike proprietary standards, users are not reliant on a private organization to license use and provide support. *See also*: nonproprietary; open architecture; open-source.

operational record

n., A record that relates to the substantive activities an organization undertakes to accomplish its mission or mandate. *Syn.*: program record.

optical disk

n., A high-density, direct access, digital storage medium consisting of a specially coated disk on which data are encoded in a pattern of tiny pits burned into the surface with a laser, to be read by a device that reflects a laser beam off the pitted surface, then decoded by a microprocessor into digital signals. *See also*: compact disc; digital videodisc; non-volatile storage.

organization

n., A social system that has an unequivocal collective identity, and exact roster of members, a program of activity, and procedures for replacing members. *See also*: agency; corporate body.

original file

n., In the paper environment, a file that contains originals of documents received and drafts and/or copies of documents sent (i.e., the first complete and effective file).

original record

n., The first copy or archetype of a record; that from which another instrument is transcribed, copied, or initiated. *See also*: first manifestation of a record.

originator

n., The person assigned the electronic address where the record has been generated (i.e., from which the record is sent or where the record is compiled and kept).

originator's name

n., The name of the person from whose electronic address the record has been sent.

outgoing document

n., A documents that is sent to an external juridical and physical person in the course of the activities of the records creator, a draft or record copy of which is also set aside by the creator, usually in a recordkeeping system. *See also*: copy of outgoing document.

outgoing record

See: outgoing document

overall presentation

n., [diplomatics] An extrinsic element of documentary form concerning a record's overall information configuration; i.e., the manner in which the content is presented to the senses using text, image or sound, either alone or in combination. *See also*: form; format; specific presentation feature.

packet
n., A piece of a message transmitted over a packet-switching network. One of the key features of a packet is that it contains the destination address in addition to the data.

packet switching
n., Refers to protocols in which messages are divided into packets before they are sent. Each packet is then transmitted individually and can even follow different routes to its destination. Once all the packets forming a message arrive at the destination, they are recompiled into the original message.

parity bit
n., A bit included in a unit of digital data to detect errors in transmission.

perfect
a., [law; diplomatics] Complete, finished, without defect and enforceable.

perfect record
n., A record that is able to produce the consequences wanted by its author; perfection is conferred on a record by its form. *See also*: authentic record; authoritative record; effective record; complete record; reliable record.

permanent preservation
See: records preservation

permanent preservation system
See: records preservation system

persistent archive
n., A strategy that seeks to make the architecture of archival information systems used to preserve digital records independent of the technology used to implement them. *See also*: persistent object preservation.

persistent format
n., A data type, which may be simple or complex, that is independent of specific hardware or software, such that an object in this data type can be transferred from a source platform to an arbitrary target platform with no significant alteration of essential attributes or behaviours. *See also*: self-describing.

persistent object
n., A digital object normalized to a persistent format and encapsulated so that it is self-describing and readable by newer platforms.

persistent object preservation
n., (POP) A permanent preservation technique to ensure digital records remain accessible by making them self-describing in a way that is independent of specific hardware and software. *See also*: persistent archive.

persistent storage
See: non-volatile storage

person
n., An individual or legally defined entity who is the subject of rights and duties, and who is recognized by the juridical system as capable of or having the potential for acting legally. *See also*: addressee; author; juridical person; originator; physical person; recipient; writer.

pertinent
a., [diplomatics] The quality of a record whose content is relevant to the purpose for which it is created and/or used. With correct, precise and truthful, a component of accuracy.

physical form
n., [diplomatics] The whole of the formal attributes of the record that determine its external make-up. *See also*: intellectual form.

physical person
n., A human being, as distinguished from a juridical person, who has natural rights and duties and who has the ability to act in his or her own right in relations with other people. *Syn.*: natural person.

physical structure
n., Constituting the data storage elements in an SGML-compliant document and their content. *See also*: document type definition; logical structure.

PKC
Initialism for "public key cryptosystem."

PKE
Initialism for "public key encryption."

PKI
Initialism for "public key infrastructure."

plaintext
n., Text or other data that contains no formatting and/or is not encrypted.

planned disposition
See: disposition rule

platform
n., The specific hardware architecture of a computer and/or its operating system; usually for a model or entire family of computers.

platform-independent
See: cross-platform

platform-neutral
See: cross-platform

policy
n., A formal statement of direction or guidance as to how an organization will carry out its mandate, functions or activities, motivated by determined interests or programs. *See also*: best practice; guideline; rule; standard.

POP
Initialism for "persistent object preservation."

potential record
n., A document that is a record in becoming. If the creator treats it is as a record, associates it with entities that are undeniably records, and does so in the course of an activity and for its purpose, such an object only needs a stable content and a fixed form to materialize itself as a complete record.

preamble
n., [diplomatics] An intrinsic element of documentary form that comprises the part of the text of a document that expresses the ideal motivation of the action.

precise
a., [diplomatics] The quality of a record that strictly conforms in every detail of content and form to an established standard, guideline or convention. With correct, pertinent and truthful, a component of accuracy.

presented digital record
See: manifested digital record

preservation
n., The whole of the principles, policies, rules and strategies aimed at prolonging the existence of an object by maintaining it in a condition suitable for use, either in its original format or in a more persistent format, while leaving intact the object's intellectual form. *See also*: persistent object preservation; records preservation.

preservation emulation
See: emulation

preservation strategy
See: records preservation strategy

preservation system
See: records preservation system

preserved record
n., A record in the records preservation system that is the result of good record-making, recordkeeping and permanent preservation practices, and is available for output (upon request).

preserver
See: designated records preserver

presumption of authenticity
n., An inference as to the fact of a record's authenticity that is drawn from known facts about the manner in which that record has been created and maintained. *See also*: authenticity requirements; benchmark authenticity requirements.

primary storage
See: volatile storage

primitiveness
n., The quality of being first, of not being derived from something else. With completeness and effectiveness, a quality presented by an original record.

priority of transmission
n., Indication of the relative importance or urgency with which a record is to be transmitted.

private key
n., In a Public Key cryptosystem, that part of a key pair that is held by a logical or legal entity in an authentication system, protected by a password, and not made available to anyone else. *See also*: public key; public key cryptosystem; public key encryption; public key infrastructure.

probative
a., (of a document) Constituting evidence of a completed juridical act. *See also*: dispositive; narrative; supporting.

probative record
n., A retrospective record for which the juridical system requires a written form as evidence of an action that came into existence and was complete before being manifested in writing. With dispositive, enabling, instructive, narrative and supporting, one of six functional categories of records.

procedural context
n., The business procedure in the course of which a record is created. *See also*: administrative context; documentary context; juridical-administrative context; provenancial context; technological context.

procedure
n., In general, the body of written and unwritten rules governing the conduct of a transaction, or the formal steps undertaken in carrying out a transaction.

process

n., The series of motions, or activities in general, carried out to set oneself to work and go on towards each formal step of a procedure.

profile

n., An analysis representing the extent to which an entity exhibits various characteristics.

program record

See: operational record

property sheets

n., A type of window or dialogue box provided in many software applications that lists the attributes or settings of an object such as a file, application or hardware device. A property sheet presents the user with a tabbed, index-card-like selection of property pages, each of which features standard, dialogue-style controls for customizing parameters.

proprietary

a., In reference to hardware technology, software applications and/or file formats, the state of being privately owned and controlled. A proprietary design or technique...implies that the company has not divulged specifications that would allow other companies to duplicate the product. *See also*: open architecture; open-source; open-standard products and formats.

prospective record

n., A record that guides what to do and/or how to do it. A prospective record can enable (enabling record) or inform (instructive record) interactions, experiences or dynamic processes. *See also*: retrospective record.

protocol

n., [record] *See*: protocol register

n., [documentary form] The initial section of a document, usually containing the identification of the persons concurring to its formation and of its temporal, geographical and administrative context.

n., [computing] A set of formal rules describing how to transmit data, especially across a network. Low level protocols define the electrical and physical standards to be observed, bit- and byte-ordering and the transmission and error detection and correction of the bitstream. High level protocols deal with the data formatting, including the syntax of messages, the terminal to computer dialogue, character sets, sequencing of messages, etc.

protocol register

n., A type of register that records the identifying attributes of incoming, outgoing, and/or internal records, specifying the action taken.

provenance

n., The relationships between records and the organizations or individuals that created, accumulated and/or maintained and used them in the conduct of personal or corporate activity. *See also*: provenancial context.

provenancial context

n., The creating body, its mandate, structure and functions. *See also*: administrative context; documentary context; juridical-administrative context; procedural context; provenance; technological context.

pseudo-original

n., A copy of a record in which the maker of the copy tries to imitate perfectly the original in order to deceive. *See also*: authentic copy; conformed copy; copy in form of original; imitative copy; simple copy.

public accountability

n., ‘Right to know,’ a right to receive openly declared facts that may lead to public debate by the citizens and their elected representatives.

public key

n., In a public key cryptosystem, that key of a user’s key pair that is publicly known. *See also*: private key; public key cryptosystem; public key encryption; public key infrastructure.

public key cryptosystem

n., (PKC) A cryptographic system that uses two keys: a public key known to everyone and a private or secret key known only to the recipient of the message.

public key encryption

n., (PKE) The use of two keys—a public key known to everyone and a private or secret key known only to the recipient of the message—to encrypt and decrypt information transmitted between two parties. *See also*: public key cryptosystem; public key infrastructure.

public key infrastructure

n., (PKI) The underlying systems and processes necessary to support the trustworthiness and wide-scale use of public key encryption to authenticate individuals in a digital environment, especially over the Internet.

publication

n., Recorded information that is intended for communication and/or dissemination to the public at large.

qualification of signature

n., [diplomatics] An intrinsic element of documentary form that comprises the mention of the title and capacity of the signer, usually accompanying the attestation.

RDF

Initialism for “resource description framework.”

received document

n., A document transmitted to a creator from an external juridical or physical person. *See also*: made documents; received records.

received record

n., A received document declared a record and set aside for action or reference, usually in a recordkeeping system. *See also*: internal record; made record.

recipient

n., Person(s) to whom the record is copied for information purposes. *Syn.*: receiver.

recipient’s name

n., The name of the office or individual receiving the record.

reconstitute

v., To link and assemble the stored digital component(s) of a document to enable the document to be reproduced and manifested in authentic form. *See also*: render; reproduce digital records.

record

n., A document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference. *Syn.*: archival document. *See also*: accession record; accessioned records; acquired records; active record; aggregated records; analogue record; archives [records]; arranged records; audiovisual record; aural record; authentic copy; authentic record; authenticated record; authoritative record;

classified record; complete record; conformed copy; copy in form of original; created record; declared record; described records; digital record; dispositive record; draft; dynamic record; effective record; electroacoustic record; electronic record; enabling record; executed record; experiential record, genuine record; graphic record; imitative copy; inactive record; incoming record; instructive record; intact record; interactive record; internal record; made record; manifested digital record; narrative record; obsolete record official record; operational record; original record; outgoing record; perfect record; pseudo-original; potential record; presented digital record; preserved record; probative record; prospective record; received record; registered record; reliable record; reproduced digital record; reproducible digital record; retrospective record; sent record; semiactive record; simple copy; stored digital record; supporting record; virtual record; written record.

record attribute

n., [diplomatics] A defining characteristic of a record or of a record element (e.g., the name of the author).

record characteristic

n., [diplomatics] A quality that belongs to all records, such as a fixed documentary form, a stable content, an archival bond with other records either inside or outside the system, and an identifiable context.

record creation

n., The first phase of a record's lifecycle in which a record is made or received and then set aside for action or reference, usually in a recordkeeping system. *See also*: create records; record-making.

record element

n., [diplomatics] A constituent part of a record's documentary form; an element is a formal expression visible on the face of the record (e.g., a signature).

record identity

n., The distinct character of a record, identifiable through the attributes that uniquely characterize it and distinguish it from other records.

record integrity

n., A record's wholeness and soundness. The quality or state of being complete and uncorrupted.

record profile

n., An electronic form designed to contain the attributes of the record that attest to its identity and integrity, and which is generated when a user tries to send or to close an electronic record and remains inextricably linked to the record for the entire period of its existence.

record version

n., One of two or more forms in which a record is issued.

recordkeeping

v., The whole of the principles, policies, rules and strategies employed by the creator that establishes and maintains administrative, intellectual and physical control on its records. *See also*: record-making.

recordkeeping access privileges

n., The authority to annotate, read, retrieve, transfer and/or destroy records in the record-keeping system, granted to officers and employees of the creator.

recordkeeping metadata schemes
n., Lists of all necessary metadata to be recorded to ensure the identification and integrity of records maintained in the recordkeeping system.
recordkeeping office
n., The office given the formal competence for designing, implementing and maintaining the creator's trusted recordkeeping system.
recordkeeping system
n., A set of rules governing the storage, use, maintenance and disposition of records and/or information about records, and the tools and mechanisms used to implement these rules. *See also*: trusted recordkeeping system.
record-making
v., The whole of the principles, policies, rules and strategies that controls the process of creating records from made or received documents. *See also*: create records; record creation; recordkeeping.
record-making access privileges
n., The authority to compile, annotate, read, retrieve, transfer and/or destroy records in the record-making system, granted to officers and employees of the creator.
record-making metadata schemes
n., Lists of all necessary record-making metadata to be recorded to ensure the reliability, accuracy, identification and integrity of records created in the record-making system.
record-making system
n., A set of rules governing the making of records, and the tools and mechanisms used to implement these rules.
records aggregation
n., A natural accumulation of an interrelated group of records, such as a file, dossier, series or fonds, which results from the way in which a records creator carries out its activities or functions. *Syn.*: aggregated records. *See also*: archival aggregation; arranged records.
records continuum
n., A model of archival science that emphasizes overlapping characteristics of recordkeeping, evidence, transaction, and the identity of the creator. *See also*: records lifecycle.
records creator
n., The physical or juridical person who makes, receives or accumulates records by reason of its mandate/mission, functions or activities and who generates the highest-level aggregation in which the records belong (that is, the fonds). *Syn.*: creator. *See also*: controlling agency; custody; records manager.
records forms
n., Specifications of the documentary forms for the various types of records of the creator.
records keeper
See: records manager
records lifecycle
n., A model of records management and archival science that characterizes the life span of a record as comprising eight sequential stages: creation or receipt; classification; maintenance and use; disposition through destruction or transfer to an archival institution or agency; description in archival finding aids; preservation; reference and use. *See also*: active records; inactive records; records continuum; semiactive records.

records management

v., The systematic design, implementation, maintenance and administrative control of a framework for the making and keeping of records by a records manager (trusted records officer) to ensure efficiency and economy in their creation, use, handling, control, maintenance and disposition.

n. The field of management responsible for the efficient and systematic control of the creation, receipt, maintenance, use and disposition of records, including processes for capturing and maintaining evidence of and information about business activities in the form of records.

records management function

n., The whole of the activities of a creator aimed at the creation, use and maintenance of records to meet its administrative, programmatic, legal, financial and historical needs and responsibilities.

records manager

n., The person responsible for the management of active and semiactive records of a creator. The role of a records manager should be that of a trusted records officer. *Syn.*: records keeper; records officer. *See also*: archivist; designated records preserver; records creator.

records officer

See: records manager; trusted records officer

records preservation

n., The whole of the principles, policies, rules and strategies that controls the physical and technological stabilization and protection of the intellectual form of acquired records intended for their continuing, enduring, stable, lasting, uninterrupted and unbroken chain of preservation, without a foreseeable end. *Syn.*: archival preservation; permanent preservation. *See also*: digital preservation; persistent object preservation; preservation.

records preservation strategy

n., A coherent set of objectives and methods for protecting and maintaining (i.e., safeguarding authenticity and ensuring accessibility of) digital components and related information of acquired records over time, and for reproducing the related authentic records and/or archival aggregations. *Syn.*: preservation strategy; permanent preservation strategy. *See also*: backward compatibility; conversion; data restoration; emulation; encapsulation; maintenance strategy; normalization; persistent object preservation; software re-engineering; technology preservation.

records preservation system

n., A set of rules governing the permanent intellectual and physical maintenance of acquired records and the tools and mechanisms used to implement these rules. *Syn.*: archival preservation system; preservation system; permanent preservation system. *See also*: persistent archive; trusted preservation system.

records retention schedule

See: retention schedule

records series

n., Dossiers, file units or individual documents that are arranged in accordance with a classification or filing system or that are maintained as a unit because they result from the same accumulation or filing process, the same function or the same activity, and that have a particular form or because of some other relationship arising out of their creation, receipt or use. *Syn.*: series. *See also*: dossier; file; fonds; item.

records system
n., A set of rules governing record-making and recordkeeping, as controlled by the creator's records management function, and the tools and mechanisms used to implement these rules.

refreshing
v., The process of copying the digital content from one digital medium to another (includes copying to the same kind of medium). *See also*: conversion; migration; transformative migration.

refreshing of records
v., The process of refreshing digital records in the usual and ordinary course of business to ensure their continued accessibility as their storage medium becomes obsolete or degrades over time, while leaving intact their intellectual form. *See also*: conversion of records; migration of records; transformative migration of records.

registered record
n., A record that has been assigned a registration number and for which all the data necessary to identify the persons and acts involved and the documentary context of the record are recorded within a protocol register.

registration number
n., A consecutive number added to each incoming or outgoing record in the protocol register, which connects it to previous and subsequent records made or received by the creator.

registration scheme
n., A method for assigning a unique identifier to each record.

registry system
n., A system controlling the creation, maintenance, and use of current and semicurrent records through the use of formal registers, lists and indexes.

reliability
n., The trustworthiness of a record as a statement of fact. It exists when a record can stand for the fact it is about, and is established by examining the completeness of the record's form and the amount of control exercised on the process of its creation.

reliable record
n., A record capable of standing for the facts to which it attests. *See also*: authentic record; authoritative record; complete record; effective record; perfect record.

render
v., To draw a real-world object as it actually appears. *See also*: manifest; reconstitute.

reproduce
v., To make a copy.

reproduced digital record
n., An authentic representation or other version of a digital record reconstituted from its digital component(s). *See also*: authentic copy; copy in form of original; imitative copy; manifested digital record; simple copy.

reproducible digital record
n., The digital component(s) of a record together with the technical information or software necessary to reproduce and manifest it from the digital component(s).

reproduction
n., The process of generating a copy.

repurposing

n., The process of taking content from one medium (such as from a book, a newspaper, TV, or radio) and repackaging it for use in another medium (such as one the Web).

requirement

n., A constraint, demand, necessity, need, or parameter that must be met or satisfied, usually within a certain timeframe or as a prerequisite. *See also*: authenticity requirement; baseline authenticity requirements; benchmark authenticity requirements; compliant.

resource description framework

n., (RDF) An XML-based language for representing information about resources in the World Wide Web. It is particularly intended for representing metadata about Web resources, such as the title, author, and modification date of a Web page, copyright and licensing information about a Web document, or the availability schedule for some shared resource.

retention schedule

n., A document providing description of records series and/or classes and specifying their authorized dispositions.

retrieval system

n., A set of rules governing searching and finding records in recordkeeping and records preservation systems, and the tools and mechanisms used to implement these rules.

retrospective record

n., A record that is a means of remembering what was done. *See also*: narrative record; probative record; supporting record; prospective record.

rule

n., An authoritative statement of what to do or not to do in a specific situation, issued by a competent person. *See also*: best practice; directive; guideline; policy; standard.

salutation

n., [diplomatics] An intrinsic element of documentary form comprising a form of greeting that appears only in letters, usually following the inscription.

save

v., To affix a digital object in non-volatile storage on a digital medium.

schema

n., A structured framework or plan.

schema document

n., An SGML-compliant document that defines the structure and contents of other SGML-compliant documents, in a similar manner to a Document Type Definition (DTD). *Syn.*: document schema.

scheme

n., A group of independent but interrelated elements comprising a unified whole.

science

n., The body of knowledge comprising measurable or verifiable facts acquired through application of the scientific method, and generalized into scientific laws or principles.

scientific method

n., A rigorous, systematic approach, designed to eliminate bias and other subjective influences in the search, identification, and measurement or validation of facts and cause-effect relationships, and from which scientific laws may be deduced.

script

n., [diplomatics] An extrinsic element of documentary form that comprises the characteristics of a document's writing such as the layout of the writing with respect to the physical form of the document, the presence of different hands or types of writing in the same document, the correspondence between paragraphs and conceptual sections of the text, type of punctuation, abbreviations, initialisms, ink, erasures, corrections, etc.

seal

n., A piece of wax, lead or other material upon which an impression has been made and which is attached to a document or applied to the face thereof. Originally serving as a means of authentication of the author of a record and of the record itself. *See also*: electronic seal.

secondary storage

See: non-volatile storage

secretarial note

n., [diplomatics] An intrinsic element of documentary form that comprises any of several types of clerical notes that might appear on a document, such as the initials of the typist, the mention of enclosures, or an indication that the document is copied to other persons.

selection system

n., A set of rules governing the appraisal of records, the monitoring of appraisal decisions, the monitoring of the performance of the selection system, and the tools and mechanisms used to implement these rules.

self-authenticating

v., Authentication without extrinsic evidence of truth or genuineness. In federal courts, certain writings, such as notarized documents and certified copies of public records, may be admitted into evidence by self-authentication.

self-describing

n., An object whose data structure, form, or layout provides both definitions and values for the data or formats of the object. A self-describing entity can be evaluated, with all its elements and forms understood, without the need of external references. *See also*: persistent format; persistent object.

semantic schema

n., The representation of a vocabulary in a particular machine-processable form, such as an RDF or relational-database schema.

semantic rule

n., A rule that governs the meanings or interpretations of symbols or elements within an object. *See also*: syntactic rule.

semiactive record

n., A record which is no longer needed for the purpose of carrying out the action for which it was created, but which is needed by the records creator for reference. *Syn.*: semicurrent record. *See also*: active record; inactive record; records lifecycle.

semicurrent record

See: semiactive record

sent document

See: outgoing document

sent record

See: outgoing document

serialize

v., To save a digital object onto a storage medium in a location-independent way so that it can be transmitted or stored elsewhere.

series

See: records series

set aside

v., To declare a record and retain it for future reference or use, usually in a recordkeeping system.

SGML

Initialism for "Standard Generalized Markup Language."

SGML document

n., A digital document encoded using Standard Generalized Markup Language (SGML) in conformance with the syntactic rules described in a Document Type Definition (DTD) or a schema document. *See also*: HTML document; XML document.

SGML-compliant document

n., A digital document encoded using Standard Generalized Markup Language (SGML) or any one of its derivative markup languages, such as Hypertext Markup Language (HTML) and eXtensible Markup Language (XML). *See also*: HTML document; SGML document; XML document.

signature

n., The name or special mark of a person, affixed by the person's hand or by its authorized agent on a document for the purpose of taking responsibility for, approving, or validating all or part of its content.

simple copy

n., A copy that only reproduces the content of a record. *See also*: authentic copy; conformed copy; copy in form of original; imitative copy; pseudo-copy.

software re-engineering

*n.,*The systematic transformation of existing software or systems into a new form to realize quality improvements in operation, system capability, functionality, interoperability, performance or evolvability. *See also*: backward compatibility; conversion; preservation strategy; transformative migration.

special sign

n., [diplomatics] An extrinsic element of documentary form that comprises a symbol that identifies one or more of the persons involved in the compilation, receipt or execution of a record.

specific presentation feature

n., [diplomatics] An extrinsic element of documentary form concerning specific aspects of the record's formal presentation that are necessary for it to achieve the purpose for which it was created, such as special layouts, hyperlinks, deliberately employed type fonts or colours, image resolutions, audio sampling rates, etc. *See also*: overall presentation.

specification

n., A detailed description of features and/or functions in the design of an entity or system.

stable

a., With reference to content, either unchangeable or changeable according to fixed rules, that is, endowed with bounded variability. *See also*: fixed form; fixity.

standard
n., The complex of established norms aiming to make the characteristic of a product, process, or service uniform within or across a sector, a country, or a system. *See also*: best practice; de facto standard; de jure standard; directive; guideline; policy; protocol [computing]; rule.

standard generalized markup language
n., (SGML) An ISO standard text-formatting language for defining descriptions of the structure and content of different types of digital documents. *See also*: hypertext markup language; extensible markup language.

standards
n., Sets of rules or guidelines co-operatively adhered to by peer entities.

state of transmission
See: status of transmission

status of transmission
n., The degree of perfection of a record; that is, whether a record is a draft, an original or a copy.

stored digital component
n., A digital object that is placed in a storage system on a digital medium and is treated and managed as a digital component.

stored digital document
n., A digital document that is placed in a storage system on a digital medium and is treated and managed as a document.

stored digital object
n., A digital object that is placed in a storage system on a digital medium.

stored digital record
n., A stored digital document that is treated and managed as a record.

strategy
n., The complex of practical means formally articulated by an entity for reaching a specific purpose, that is, a plan or a road map for implementing policies.

subject
n., [diplomatics] An intrinsic element of documentary form that comprises a symbol that identifies one or more of the persons involved in the compilation, receipt, or execution of a record.

subject line
n., Location of statement signifying what a document is about such as specific topics, functions, or activities.

subscription
v., The act of signing one's name on a document; the signature so affixed.

superscription
n., [diplomatics] An intrinsic element of documentary form that comprises the mention of the name of the author of the document and/or the action and which may take the form of an entitling.

supporting
a., (of a document) Constituting evidence of a juridically relevant activity. *See also*: dispositive; narrative; probative.

supporting record
n., A retrospective record constituting written evidence of an activity that does not result in a juridical act, but is itself juridically relevent. With dispositive, narrative, probative and prospective, one of five functional categories of records.

symbol
n., A character, image, mark, shape, characteristic, or thing used to represent or denote something else by association, convention, or unintended resemblance.

syntactic rule
n., A rule that governs the ways symbols or elements within an object can be arranged and used. *See also*: semantic rule.

system
n., An organized and integrated set of detailed methods, policies, procedures, resources, routines, rules and tools established or formulated to carry out a specific activity, perform a duty, or solve a problem.

tag
See: mark-up; markup tag

technological context
n., The characteristics of the hardware, software, and other components of an electronic computing system in which records are created.

technology preservation
n.,. A digital records preservation strategy that involves maintaining the original software and hardware platforms with which the records were created or last manifested in authentic form.

technology-dependent authentication
n.,. The use of technological mechanisms, such as digital signatures or other cryptographic techniques, to authenticate records; technology-independent authentication.

technology-independent authentication
n., The authentication of records based on the use of administrative procedures to establish a presumption of authenticity or, if necessary, a verification of authenticity, especially through comparison of the evidence compiled about a record's identity and integrity and the procedural controls exercised over its creation, use, maintenance and/or preservation with the requirements for authentic records. *See also*: baseline requirements; benchmark requirements; technology-dependent authentication.

terms and conditions of transfer
n., Formal instruments that identify in archival and technological terms digital records to be transferred, together with relevant documentation, and that identify the medium and format of transfers, when the transfers will occur, and the parties to the transfers.

text
n., A collection of words, numbers, or symbols that conveys meaning as language.
n., [documentary form] The central section of a document, which contains the action, including the considerations and circumstances which gave origin to it, and the conditions related to its accomplishment.

third-party intellectual property rights
n., Rights belonging to a party other than the author or the addressee of a record.

timestamp
n., An attestation by a trusted third party that a record was received at a particular point in time.

title
n., [diplomatics] An intrinsic element of documentary form comprising an indication of the action, matter of subject of the record, usually under the entitling or in its place.

tools
n., Information, technology and other equipment and supplies used to manage the lifecycle of records.

topical date
n., The place of the compilation of a record, included in the record by its author. *See also*: chronological date.

transaction
n., An act or several interconnected acts in which more than one person is involved and by which the relations of those persons are altered.

transformative migration
n., The process of converting or upgrading digital objects or systems to a newer generation of hardware and/or software computer technology. *See also*: conversion; migration; refreshing; upgrade [*v.*].

transformative migration of records
n., The process of converting records in the usual and ordinary course of business (otherwise the activity is not migration but creation) to maintain their compatibility with a newer generation of hardware and/or software computer technology, while leaving intact their intellectual form. *See also*: conversion of records; migration of records; refreshing.

transmission
v., The moving of a record across space (from a person or organization to another, or from a system to another), or through time.

trusted custodian
n., A preserver who can demonstrate that it has no reason to alter the preserved records or allow others to alter them and is capable of implementing all of the requirements for the preservation of authentic copies of records. *See also*: archivist; authenticity requirement; baseline authenticity requirements; designated records preserver; records manager; trusted records officer; trusted third party.

trusted preservation system
n., The whole of the rules that control the preservation and use of the records of the creator and provide a circumstantial probability of the authenticity of the records, and the tools and mechanisms used to implement those rules. *See also:* authenticity requirement; baseline authenticity requirements.

trusted recordkeeping system
n., The whole of the rules that control the creation, maintenance use and disposition of the records of the creator and provide a circumstantial probability of the authenticity of the records, and the tools and mechanisms used to implement those rules. *See also:* authenticity requirement; benchmark authenticity requirements.

trusted records officer
n., An individual or a unit within the creating organization who is responsible for keeping and managing the creator's records, who has no reason to alter the creator's records or allow others to alter them and who is capable of implementing all of the requirements for authentic records. *See also*: archivist; authenticity requirement; benchmark authenticity requirements; designated records preserver; records manager; trusted custodian; trusted third party.

trusted third party

n., Any outsider or person not a party to the fact or act nor immediately concerned with it. *See also*: trusted custodian; trusted records officer.

trustworthiness

n., The accuracy, reliability and authenticity of a record.

truthful

a., [diplomatics] The quality of a record whose content is in accordance with the actual state of affairs. With correct, pertinent and precise, a component of accuracy.

UDF

Initialism for "uniform disk format."

unbroken custody

n., A traceable and uninterrupted line of care, control and usually possession of a body of records from creation to preservation that can serve as a means of protecting the authenticity of the record. *See also*: chain of preservation.

unicode

n., A universal 16-bit (two byte) standard character set for representing all scripts in active modern use as plain text in computer processing. *See also*: ASCII; UTF-8.

uniform resource locator

n., (URL) A standard way of [uniquely] specifying the location of an object, typically a Web page, on the Internet; the form of address used on the World-Wide Web. *Syn.*: universal resource locator.

universal character set transformation format 8

n., (UTF-8) An ASCII-compatible scheme for encoding Unicode values in sets of eight bits.

universal disk format

n., A universal, vendor-independent file system standard (ISO 13346) for storing data on optical media; designed for data interchange and portability, allowing an operating system to read, write and modify data stored on optical media that were created by another operating system.

universal resource locator

See: uniform resource locator

upgrade

n., A new or better version of some hardware or software computer technology.

v., To develop or install a new or better version of some hardware or software computer technology. *See also*: conversion; transformative migration.

URL

Initialism for "uniform resource locator."

UTF-8

Initialism for "universal character set transformation format 8."

verification of authenticity

n., The act or process of establishing a correspondence of known facts about the record itself and the various contexts in which it has been created and maintained with the proposed fact of the record's authenticity.

version

n., One of several variations of an intellectual work, possibly created for a purpose or use other than the one originally intended.

virtual record

n., A digital document perceived as existing by the user, but not existing in the system as seen.

vital record code

n., The indication of the degree of importance of a record to continue the activity for which it was created or the business of the person/office that created it.

volatile storage

n., Computer memory that requires electrical power and, in some cases periodic refreshment (e.g., DRAM), to maintain its stored content. *Syn.*: primary storage. *Opp.*: non-volatile storage.

WAN

Initialism for "wide area network."

wide area network

n., (WAN) A data network usually constructed over distances greater than one kilometre. *See also*: local area network (LAN) and metropolitan area network (MAN).

word

n., [computing] A fundamental unit of storage in a computer. The size of a word in a particular computer architecture is one of its chief distinguishing characteristics. *See also*: byte.

work

n., A distinct expression of human thought or emotion made in language, signs, symbols, numerals, images, or some other form, for purposes of communication and remembrance. *See also*: document; record.

wrapper

n., A data structure or software that encapsulates ("wraps around") other data or software objects, appends code or other software for the purposes of improving user convenience, hardware or software compatibility, or enhancing data security, transmission or storage. *See also*: emulation; encapsulation.

wrapper format

n., A specified wrapper structure for encapsulating multiple bitstreams into a single file.

writer

n., Person having the authority and capacity to articulate the content of the record.

writer's name

n., The designation (name) of the person competent for the articulation of the content of the record.

written record

n., A document created by a physical or juridical person in the course of practical activity that is produced on a medium (paper, magnetic tape, disc, plate, etc.) by means of a writing instrument (pen, pencil, typing machine, printer, etc.) or of an apparatus for fixing data, images and/or voices.

XML

Initialism for "eXtensible Markup Language."

XML document

n., An SGML-compliant digital document encoded using eXtensible Markup Language (XML) in conformance with the syntactic rules described in a Document Type Definition (DTD) or a schema document. *See also*: HTML document; SGML document.