



Aslib Proceedings

Emerald Article: Long-term digital information preservation: challenges in Latin America

Juan Voutssas

Article information:

To cite this document:

Juan Voutssas, (2012), "Long-term digital information preservation: challenges in Latin America", Aslib Proceedings, Vol. 64 Iss: 1 pp. 83 - 96

Permanent link to this document:

<http://dx.doi.org/10.1108/00012531211196729>

Downloaded on: 26-06-2012

References: This document contains references to 19 other documents

To copy this document: permissions@emeraldinsight.com

This document has been downloaded 271 times since 2012. *

Users who downloaded this Article also downloaded: *

Gobinda Chowdhury, (2010), "From digital libraries to digital preservation research: the importance of users and context", Journal of Documentation, Vol. 66 Iss: 2 pp. 207 - 223

<http://dx.doi.org/10.1108/00220411011023625>

Aryati Bakri, Peter Willett, (2011), "Computer science research in Malaysia: a bibliometric analysis", Aslib Proceedings, Vol. 63 Iss: 2 pp. 321 - 335

<http://dx.doi.org/10.1108/00012531111135727>

Andrew M. Cox, Megan K. Blake, (2011), "Information and food blogging as serious leisure", Aslib Proceedings, Vol. 63 Iss: 2 pp. 204 - 220

<http://dx.doi.org/10.1108/00012531111135664>

Access to this document was granted through an Emerald subscription provided by UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

For Authors:

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service.

Information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

With over forty years' experience, Emerald Group Publishing is a leading independent publisher of global research with impact in business, society, public policy and education. In total, Emerald publishes over 275 journals and more than 130 book series, as well as an extensive range of online products and services. Emerald is both COUNTER 3 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.



Long-term digital information preservation: challenges in Latin America

Long-term digital
information
preservation

Juan Voutssas

*Centro Universitario de Investigaciones Bibliotecológicas,
Universidad Autónoma de México (UNAM), Mexico City, Mexico*

83

Abstract

Purpose – The purpose of this paper is to summarize the current situation of digital preservation and describe the additional main challenges and issues faced in Latin American countries within the framework of the factors of digital preservation.

Design/methodology/approach – This paper reviews the literature on existing digital information preservation around the world, and then focuses on Latin American countries. It proposes a framework in which to analyze the situation in developing countries, particularly in Latin America, and summarizes the issues in six preservation factors and seven principles of preservation.

Findings – The amount of digital information has increased geometrically without proper knowledge, theories, strategies, policies and proceedings to preserve it. Developed countries and regions have started to create this knowledge and know-how, as well as setting the basis to manage their digital collections within libraries, archives, and other organizations with the aim to preserve them in the long-term. The situation is radically different in Latin America where we found limited research and projects on digital preservation.

Practical implications – The problem of the long-term authentic preservation of digital document collections is not a simple technology problem; it is a much more complex one. This paper summarizes the approaches to understand it but further study is required in order to develop a practical strategy and/or plan. It also highlights the importance of raising awareness of digital preservation in Latin America in order to ensure availability in the future.

Social implications – Research, planning, initiatives and proper execution in regard to digital collections preservation is extremely limited in Latin American developing countries, thus endangering vast amounts of digital information. There is a serious risk that valuable digital information will not be preserved adequately for future generations.

Originality/value – This paper reviews the main challenges for libraries and archives to preserve digital information over the long term and summarizes the main issues to understand and to thwart this problem. It also highlights the often overlooked problem of digital preservation in developing countries, in particular Latin America.

Keywords Digital archives Long-term preservation, Digital libraries, Digital information, Information preservation

Paper type Research paper



Each renowned author from ancient times who I recover from oblivion places a new offense and a new cause of dishonor to the charge of preceding generations who, not satisfied with their own disgraceful shallowness – allowed that other fruitful minds works, the writings that their ancestors produced with great toil and endeavor, to perish through unforgivable neglect. Nothing from their own they had to hand down to those who were to come after, but it would have been enough to forgive them just not to steal posterity of its ancestral heritage (Francesco Petrarch, ca. 1350).

Introduction

One of the main issues that Information and Communications Technology (ICT) introduced in our lives in the last decades is the vast amount of information that is produced and stored nowadays by organizations or persons. Due to the low-cost, minimum space and little effort, there has been a fundamental change in the way we perceive information storage. It is extremely difficult to attempt to determine the size of this information production and storage. Based on Lyman and Varian (2000, 2003) studies, John Ganz[1] established that by the end of 2010, and solely for that year, the world would be producing almost a Zettabyte of information[2]. In another approach, a report of the University of California, San Diego affirms that the US average user consume around 34 Gigabytes per day; Bohn and Short (2009). No figures could be found for Latin America; however, even with a conservative estimate of about 10 to 20 per cent from developed countries production and consume of the total amount, it is still several Gigabytes per day.

Almost in the same proportion that production and variety of information has increased, so has storage. Storage devices have kept pace with increase in production: they are smaller in size, greater in capacity, more reliable in performance and cheaper in price. However, issues related to long-term preservation (selection, readability, trustworthiness, among others) present a formidable challenge. Digital preservation is still a partially solved problem^{3,4}. In the words of Kathleen Shearer: "Will the digital assets being created in Canada today be accessible in five, ten and 15 years? This is a question that no one can answer right now. The life span of a digital object is very short if it is not actively managed and preserved. There is a risk that much of today's data, pictures, books, websites, software – indeed anything digital – may be lost through hardware degradation, software obsolescence, or simply a lack of capacity and resources to capture the growing volume of digital information"[5].

Initially, digital preservation was generally treated as a technology problem. Previous research tended to focus on the endurance of CD's, DVD's, tapes and other storage devices, and its artificial aging and how to keep them safely[6]. There is a large amount of literature on emulation, migration, refreshing, computer transforming and other technological aspects of digital preservation^{7,8}. The overall focus was on preservation as an ICT issue and thus its solution fell on ICT staff. However over the past few years, there is a growing awareness that this is a much more complex problem, technology is only one of its constituents and that the problem is far from being solved. Thus, it involves librarians, archivists, museologists, sociologists, economists, planners, jurists, and other kinds of information preservers beside ICT staff. Accordingly, in the last years countries have begun to design digital preservation initiatives with this holistic vision (Beagrie (2003), Digital Libraries Federation (2006), Verzosa (2005)).

This modern holistic multifactor approach has been already applied in developed countries. The main initiatives and production of knowledge are mostly from North America, Europe, and Australia. In the last 15 years several approaches have been developed, known as "models" or "reference frameworks" for this subject that try to define and to establish the attributes of documents, procedures, systems, organizations, etc., involved in the management and preservation of digital materials. Each one of them emphasizes the types of materials, elements, or attributes that are considered important for the establishment of the specific approach.

For instance, in archive science, the ISO 15489:2001 standard: "Information and Documentation: Records Management: Part 1. General" consists of a guide for the file documents management of an organization, regardless of their media. This standard establishes that "[...] the documents within an archive should be authentic, reliable, complete, without alteration, and they should permit their use and access. Likewise, they should have enough metadata to define the context, content and structure and they should reflect with precision the communication, action or decision"[9].

The Open Archival Information System, usually referred to as the OAIS model, is a reference model that has been widely accepted by the digital preservation community as a key standard for digital repositories. This model establishes the way digital collections should be preserved for a community of users from the moment digital material to the creation of a retrieval package for the end user. The OAIS reference model is a high-level reference model, i.e. it is flexible enough to use in a wide variety of organizations. The OAIS model was adopted as an ISO standard in 2003 (ISO 14721:2003 OAIS)[10].

Developed jointly by the DCC – Digital Curation Centre – and DPE – Digital Preservation Europe, the Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) represents the main intellectual outcome of a period of pilot repository audits undertaken by the DCC throughout 2006 and 2007. It establishes a methodology for self-assessment, encouraging organizations to set a comprehensive self-awareness of their objectives, activities, and assets before identifying, assessing and managing the risks implicit within their organizations[11].

The International Research on Permanent Authentic Records in Electronic Systems (InterPARES) led by the University of British Columbia aims at developing the knowledge essential to the long-term preservation of authentic records created and/or maintained in digital form and providing the basis for standards, policies, strategies and plans of action capable of ensuring the longevity of such material, and the ability of its users to trust its authenticity[12]. There are many other initiatives for digital preservation in several countries, such as the NDIIPP, The National Digital Information Infrastructure and Preservation Program, run by the Library of Congress in the USA; The Digital Preservation & Records Management Programme in the UK, run by the British Library, and the National Archives; The National Preservation Initiative in Holland, to name a few. For a wider view of these initiatives, see Beagrie (2003).

Digital preservation in Latin America

We found no significant research literature or preservation initiatives in Latin American countries, only a few isolated attempts in some countries of the region: Brazil, Chile, Argentina, Mexico, but none of the size and significance from those in other more developed countries and regions. There are in the region projects and initiatives regarding digital libraries, digital archives, digital publishing, information society, and so forth, such as Scielo (Scientific Electronic Library On-Line) in Brazil[13], e-lac (e-society in Latin America and the Caribbean)[14], but these are digitization and publishing projects, not digital preservation. The closest approaches to the problem, and only regarding digital archives, are Brazil, Mexico and the brand new Colombia "TEAMS", which are sections of the *International Research on Permanent Authentic Records in Electronic Systems*, InterPARES), a world preservation research endeavor led by the University of British Columbia[15]. As an example of the lack of significant

projects, a search on the UNESCO Archives Portal using the query: “digital preservation Latin America” only retrieves a few small collections in the “Memory of the World” project but no further initiatives or relevant research were found[16]. The problem is becoming particularly acute, mostly unnoticed and thus overlooked – in all countries of this region. Most countries are actively working on digital libraries and archives but with little or no regard for the long-term digital future. A lot of the work within the region is usually done with collections in libraries and archives that are usually neglected and have minimum budgets, lack of trained professional personnel, and practically no strategic plans for constructing and preserving long-term digital document collections. With the current ease, low-cost and availability of massive storing devices the region is already creating huge amounts of document collections with little or no awareness of how to deal with their future preservation in a proper way. This is undoubtedly endangering the greatest part of those collections in the future.

As noted previously we can distinguish in the literature the already mentioned “technology approach”, typical from the 1980s and the 1990s; a “late twentieth century approach” which established that technology was solely responsible for digital materials preservation problems and therefore the answers to this predicament would be exclusively through technology. Now we have developed the current integrative multifactor vision or “twenty-first century approach”: complex components of the problem – complex components for the solution. The aim of this article is to propose a framework in which to analyze the situation in developing countries, particularly in Latin America. As preservation is no longer considered solely a technological issue, then preservation must be addressed from a holistic, multifactor approach. We summarize all those issues in six preservation factors and seven principles of preservation as a framework to analyze the current situation in Latin America in order to issue general recommendations for digital preservation projects in this region.

Defining the issues

In order to address the issues surrounding long-term digital preservation of digital information it is important that we briefly define certain terms. In this article we shall refer to digital preservation as “the specific process of maintaining digital materials during and across different generations of technology over time, irrespective where they reside; this regards the whole of the principles, strategies, policies, rules and proceedings 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”. Long-term is defined as “a portion of time, a chronological division long enough to be concerned with the impacts of changing technologies, including new operating systems, data and device formats and standards, support for old and new media, or even with the impacts of a changing user community” (InterPARES, 2006).

The factors of preservation

There are many and complex issues involved in long term digital preservation. In order to understand them better, I have gathered these issues into six groups or “factors”: cultural, technological, legal, methodological, economic and social factors. Each one of these groups represents a particular aspect. In general terms, these factors are rarely isolated and each one of them tends to affect one another.

Cultural factors have to do with the lack of awareness of large groups of persons within current societies – including planners and decision makers – about the historical value and significance of their digital documentary heritage. This in turn leads to obliviousness to perform adequate and proper keeping of those documents with a consequent loss of heritage. Although we currently consider that we live in the information age, there exists a deep cultural problem regarding preservation of digital information; a kind of current “cultural oxymoron”. Although we value digital information production, there is not enough awareness of preserving it. In 2003, a US survey carried out by the Cornell University Library found that the main menace to digital materials was the lack of policies and plans inside their institutions to carry out this task. This problem was pointed as number one, even above the technological or economical issues[17]. In developing countries, there is an increasing awareness of this issue.

This cultural factor is particularly acute within developing countries, for example in the Latin American region, where the number of organizations that do not preserve digital materials properly – or do not preserve them at all – is very high. In this context, Petrarch’s opening quote of this paper where he condemns previous generations for neglecting to value and preserve the production of others, is particularly meaningful. A crucial issue for this region is increasing awareness of the problem.

Technological factors have to do with the rapid and unstoppable shift of devices, practices, and aspects related to ICT. The challenge is not only to maintain the information bit chains, but to do so in such a way that they are readable, meaningful and therefore, useful in the future. Technological factors are mainly related to obsolescence of computers, storage devices and media; changes in operating systems, formats, programs, interfaces, reading and reproducing devices; emerging new standards and methodologies to carry out the task; lack of interoperability among elements. All of these affect in time the permanence – the quality or state of media to endure; the accessibility – the capability of digital materials to be read and reused; the trustworthiness – the compounded reliability, authenticity, and accuracy of a document, etc. Moreover, information security must also be addressed. This has to do with the relationship among threats, risks, vulnerabilities, impacts, and control measures of digital objects assets[18]. The challenge is how to keep those innumerable records from destruction or tampering due to accidents, negligence or hacking? How to assure their integrity over long periods? One of the main problems in developing countries, such as Latin American ones, is that there is a limited capacity to produce the technology and the know-how to cope with these obsolescence and security problems. In this sense, there is a technological dependence on developing countries to solve these factors and leads to a more passive role in this subject.

Legal factors around information preservation have to do with achieving the appropriate and delicate balance between protecting copyrights and confidentiality while defending rights to information access. Who is legally responsible for keeping every document collection or archive for the future? Who has the legal eligibility or competence to perform that task? Will they be able to distribute it in the future? National libraries and archives are currently overwhelmed trying to balance their responsibilities of receiving, keeping and providing access to documents and the growing restrictions to distribute them, mostly now with the electronic formats. Lawrence Lessig (2001, 2004a, b), Jonathan Zittrain (2008), and WIPO have exhaustive works about this matter.

Another important legal issue is the right to privacy. This right in this legal context is concerned with identifiable data relating to a person that is collected and stored in digital forms. The concept of privacy may differ quite a lot from country to country and therefore, laws are unequal in-regard to this issue. What is legally public in one country may not be in another; what is private or confidential at a particular moment may change over time and space (Schneier, 2008; Solove, 2007). Although developing countries are leading the way in addressing these issues, it is important that this be addressed on a global level, drawing from the participation of countries from around the world, including the Latin American region.

Methodological factors, also known as “documentary” factors are among the most often neglected factors, particularly in archives. These factors are associated with the tools and standards used for appraisal among the different materials, proper selection and disposal, logical storing and future retrieval of documents. It is already known that just descriptive metadata – author, title, keywords – is currently not enough for proper future retrieval.

According to Tim Berners-Lee – who coined the term and the proposal, we are moving towards the “Semantic Web”[19] and from there to semantic libraries and archives. Derived from the semantic Web original concept[20], a “Semantic Library/Archive” it’s “a digital library/archive whose collections are integrated by ‘functional documents’”. Such a document starts with a digital object – like those in the current digital libraries/archives, which is a ‘faithful digital reproduction’[21] i.e. a digital object that is optimally formatted and described with a view to its quality (functionality and use value), persistence (long-term access), trustworthiness, and interoperability (e.g. across platforms and software environments). Moreover, this digital document has not just a simple set of metadata attached. It has too a semantically rich set of metadata; i.e. the usual set of metadata added with hyperlinks and context metadata. These new metadata allow to link and contextualize the document in relation with other documents enhancing its reuse, search, linking, weighting, integration, data mining and interoperability with other programs who could use them. In brief, they add ‘meaning’, ‘context’ to the document in its relation to other digital documents and the user needs”.

The usual set of metadata will still be necessary, as well as structural metadata – how objects are put together; rights metadata – the conditions to use or reproduce the document and who can access it; administrative metadata – when and how it was created, acquisition information, location information, version control and differentiation between similar information objects; technical metadata – file type, format, digitization information, like compression and scaling ratios; authentication and security data, like encryption keys or digital signature procedures; juridical metadata – laws and regulations which rule the record creation; procedural metadata – the creation procedure of a record; functional metadata – interoperability, interfaces, hyperlinks and semantic relationships among documents; preservation metadata – how document must be kept to assure proper trustworthiness, permanence, and so forth, as well as migration and refreshing already performed on records. This complex array of tools and standards are complicated but indispensable factors for proper preservation and retrieval. If these factors are not taken into consideration then the technological preservation effort will be useless, regardless of complexity and cost.

Economic factors are of course particularly important. What are the costs of digitizing and preserving a document collection and how should these be calculated? The costs of preserving can be summed up as: cost of digitizing (cost of scanning and/or producing a digital original); cost of editing (to prepare, assemble, alter, adapt, refine or bring about conformity to a standard certain digital document); cost of register (to add the set of metadata pertinent to the digital object); cost of storing (cost to maintain in storage devices in or off-line a digital object for a given time) and cost of updating (cost to copy, update, refresh, convert, and reshape digital documents to fulfill new requirements). A non-direct cost is training cost. This is particularly relevant in developing countries as there are few library professionals with the necessary skills. Additionally, preservation is an on-going process. Therefore, the current, short, and long term costs and funding are important issues to deal with before and during a preservation project, in order to maintain their feasibility in the long term. It is also important to note that, however expensive a preservation project may seem, not preserving certain information may ultimately be much more costly.

Last but not least, there are social factors. The following step after securing the preservation of document collections is to think about how to assure enduring access and usability. It is important that future generations have effective and granted access to the information that we are preserving. There is no use in preserving if no one or just a few will have access to those documents. Assuming copyrights, privacy rights, and other legal issues are observed, the future challenge will be how to make this information available to many people throughout many generations. Worldwide we are already currently dealing with what is known as the “digital gap” or “digital divide”. It is important that issues related to preservation and long term access do not increase this divide, but rather contributes towards closing it. These factors are particularly relevant in Latin American countries, where social inequality has been and still is a deeply entrenched problem. It is important that we address and take into consideration our particular social issues when working towards an “information society”, in order to promote social equality of opportunities now and in the future.

We can consider these six factors as useful aids to enable the design and execution of a strategic plan for a single or several digital collections within an institution or in various organizations. The factors represent the “big picture” around the preservation problem and constitute the framework to cope with these issues on a strategic level. On a more tactical and daily level, librarians, archivists and other information professionals have developed through long time suitable issues that can be grouped in several goals or “principles” that are very useful when dealing with digital preservation.

The principles of documents in preservation planning

To summarize numerous authors throughout different information disciplines we can group the issues in seven principles to be considered as “targets” in documents within a preservation project. These seven principles are particularly functional tools in the design and practical implementation of digital preservation efforts. Briefly, the seven principles are: selection of documents due to its documentary, academic, historical value and other issues such as use, condition, property, convenience, non duplication, among others. In archival science this is called “appraisal”. Quality – which is in direct proportion with the degree that the digital document emulates an “original” and endows it with “usability”. Permanence – has to do with the concept of the future existence of the

bit stream representing a certain digital document; making it available for the “long-term”. Accessibility – has to do with the possibility that the digital document, while existing in the future, will still be accessible, i.e. it can be interpreted, read, executed and/or displayed again. As put by (Duranti, 2005) “we cannot preserve digital material as is; we only can maintain our possibility of recreate it again and again in such a way that we can prove they are authentic copies”. Availability – has to do with the possibility, that a document can be accessed by a particular user, in a timely and suitable way and form. Availability determines who, how, where and when will be able to access the document; because of this, availability is also related with confidentiality and privacy. Functionality – this has to do with the intrinsic characteristics of a document, which will help the ICT interfaces to search, find, and link a document. It is related to the amount, variety, and quality of the intrinsic metadata of the document, but also with its semantic issues, interoperability factors, construction standards, context completeness, metadata relationships, and other factors. This principle is the basis of current theories about future “semantic libraries” and “semantic archives”. Trustworthiness – a fundamental principle in archival science. It is given when a digital record meets previously established requirements of a certain organization, which are the specifications of the elements of form and context that need to be preserved. It can be defined as “the degree of accuracy, reliability and authenticity of a record”. When a document has trustworthiness, it will have acceptability; i.e. it will be accepted for those who examine the record assuming it is trustworthy and thus it will be useful for the purposes that the record was created and also as a source of reliable information.

As it has been shown, the previously described six factors and seven principles give us the strategic and the tactical panorama about most of the issues that we must have in mind to effectively cope with preservation of digital collections, in a wide variety of contexts and environments. They must be studied and weighed carefully in detail in order to fully understand, combine and master them and thereafter become able and skilled enough to deal with projects of document preservation, despite their size and complexity. These are the set of aids that can be useful towards developing good strategies, plans, and projects and the best way to achieve an integral solution plan.

Analysis for digital preservation in Latin America

The overgrowing production, and storage of digital materials, and the everyday obsolescence of digital equipment, applications, formats, and standards, along with the lack of awareness of large groups of persons within current societies – including planners and decision makers – about the historical value, and significance of their digital documentary heritage is endangering the future existence of vast amounts of data and information. The lack of significant preservation projects in the Latin American region reflects clearly how acute this situation is there. In many cases, preservation is undertaken by individuals or organizations without knowledge, method and standards maintaining instead of preserving the digital materials with little to none degree of authenticity and reliability. Those materials, – should they exist in the future, will probably be useless because of this. The problem will be acute in the short term – ten or 20 years – but it can be catastrophic in the long term – 50 to several hundred years – if proper measures are not taken now to thwart the possible effects of poor preservation of those digital documents.

Therefore, it is important to emphasize that special efforts should be made by Latin American planners, government officers, library and archive directors, universities, research and other academic institutions in order to revert this trend developing suitable strategies and plans within the countries of the region, or even better, developing regional projects and/or agreements about how to start local research on this subject, how to build and propose significant preservation initiatives, how to prepare professional and skilled staff, how to develop and publish regional knowledge about this matter finding and exchanging local solutions and how to teach government, private organizations and even persons the way to preserve properly in order to revert these harmful effects to their collections and their documentary heritage. The projects and solutions developed in other regions and countries of the world are useful guidelines, but they cannot just be taken as an “out of the box” solution and implemented “as is” in the countries of this region. There are significant cultural, legal, and economic differences that must be adjusted or modified before their construction.

There are many ways to carry out such an endeavor. Using the six factors as a reference framework we can write a brief outline of the steps that could be carried out in the region and its countries as a general strategy to cope with this problem:

- (1) In first place, special attention must be paid to the “cultural factor”: the lack of awareness and concern of large groups of persons within current societies – including planners and decision makers – about the historical significance of their digital documentary heritage and the risk they face. A *sine qua non* condition is to develop a strategy to start destroying obliviousness in the region and in each country to prevent a consequent loss of heritage. A basic strategy has to be designed to cope with the digital information preservation as an integral problem and solution. It is fundamental that we design and develop specific suitable policies and plans within the countries of the region, or design and develop regional and national projects and/or agreements to perform adequate and proper keeping of those documents.
- (2) Certainly, the technological issues: obsolescence of computers, storage devices and media; changes in operating systems, formats, programs, interfaces, reading and reproducing devices; emerging of new standards and methodologies; lack of interoperability among elements, etc. should be taken into consideration. But, additionally social issues must be included to assure future accessibility from large amounts of people. Accordingly, a careful and adequate balance of this item must be done with the legal issues, the restrictions due to copyrights, confidentiality, and privacy of information.
- (3) The current, short and long term costs and funding are important issues to deal with before any preservation project, in order to maintain their feasibility in the long term. It is also important to remember that, despite how expensive a preservation project may seem – not preserving certain information may ultimately be much more costly.
- (4) Documentary issues: the proper structures and suitable metadata required for each type of document in each type of library, archive, or collection. The seven principles are highly valuable assets while regarding this issue. Make up-to-date inventories of all collections and assortment of materials to deal with.
- (5) Build and propose significant and sustainable preservation initiatives, aiming to:

- Create huge assets of digital documents, which are properly managed and preserved, preferably with regional and countrywide scopes.
- Prepare professional and skilled staff in this field.
- Develop regional research and knowledge about this matter finding local problems and solutions. Ensure that discoveries and adjustments are taken one step further and used to develop new theories, approaches and models; new skills, capabilities, competences, etc. Promote the systematization and publication of information. Latin American knowledge is eagerly needed in this field.
- Involve in the projects and solutions a wide range of private and public organizations and persons: librarians, archivists, museologists, ICT and other information professionals, sociologists, economists, planners, jurists, etc., and of course the research community.
- Use the standards, models, projects, methodologies and solutions developed in other regions and countries of the world as useful guidelines, but adjust them to our own cultural, legal and economic differences before their implementation.

Conclusions

Now that ICT, Internet, document digitizing, massive document storage and other technological factors are firmly established in all kind of environments in most countries, it is important that government officers, planners, decision makers; libraries, archives and museums directors; sociologists, economists, jurists, universities academic researchers and others participate in ambitious preservation projects within the Latin American region.

On a strategic level, the big framework is given by the “factors” involved in long term digital preservation. They are several and they are complex: cultural, technological, legal, methodological, economic, and social factors. Each one of these groups represents a particular aspect rarely isolated, each one of them tending to affect one another. Understanding the whole of those factors and not only the technological one enables to deal with a “twenty-first century approach”: complex components of the problem – complex components for the solution.

On a tactical level, librarians, archivists, ICT and other information professionals and the research community in other regions of the world have developed handy techniques that can be summed up into several principles that are very helpful while dealing with document preservation projects: selection, quality, permanence, accessibility, availability, functionality and trustworthiness. They give us a wide panorama of mostly all the elements that we must have in mind as “targets” on a daily basis to cope with preservation of digital collections, in a proper and professional way and in a wide variety of contexts and environments.

There are many approaches and standards to use as a basis to depart from; many ways to carry out such an endeavor. But the projects and solutions developed in other regions and countries of the world cannot just be taken as “silver bullets” solutions and implemented as they are in the countries of this region. There are significant cultural, legal, and economic differences that must be adjusted or modified before their construction. Using the six factors as a reference framework, we outlined a brief list of

the steps that could be carried out in the region and its countries as a general strategy to thwart this problem.

Special efforts should be made by the Latin American region in order to revert the negative trend by developing suitable strategies and plans within the countries of the region – or even better – developing regional projects and agreements fostering local research on this subject, building significant preservation initiatives, preparing professional, and skilled staff and sharing detailed description of results from these experiences. This will enable the region and its countries to establish the impact that ICT has had on organizations related with digital materials and how particular solutions have been designed and implemented in different environments, in order to create a “regional know-how” about digital preservation. It is very important that countries, organizations, and persons, through government, private and academic institutions, endeavor in these kind of projects, focus on and describe experiences, research, build and share practical implementations of real life situations in all kind of governmental and private environments in developing countries, so we can exchange this knowledge and know-how and learn from each other in a closer context and thus, adapt, recreate and implement better local solutions in different environments within the region.

All web references have been verified available on February 14, 2011.

Notes

1. Ganz, J. (2007), “A forecast of worldwide information growth through 2010”. International Data Corporation (IDC). March, available at: www.emc.com/collateral/analyst-reports/expanding-digital-idc-white-paper.pdf
2. 1 Megabyte = 1'000,000 = 10^6 bytes or characters. 1 Gigabyte = 1000 Megabytes = 1'000,000,000 = 10^9 bytes or characters. 1 Terabyte = 1000 Gigabytes = 1'000,000,000,000 = 10^{12} bytes or characters. 1 Petabyte = 1000 Terabytes = 1'000,000 Gigabytes = 1'000,000,000,000,000 = 10^{15} bytes or characters. 1 Exabyte = 1000 Petabytes = 1'000,000 Terabytes = 1'000,000,000,000,000,000 = 10^{18} bytes or characters. 1 Zettabyte = 1000 Exabytes = 1'000,000 Petabytes = 1'000,000,000,000,000,000,000 = 10^{21} bytes or characters (These are approximate quantities, since bytes are actually measured in powers of “2”, not in powers of “10”).
3. Shaw, J. (2010) Digital preservation: an unsolved problem, Harvard Magazine, May-June, available at: <http://harvardmagazine.com/2010/05/digital-preservation-an-unsolved-problem>
4. SNIA Data Management Forum [The Storage Networking Industry Association]. (2007) “100 year archive management survey”, available at: www.snia.org/forums/dmf/knowledge/100YrATF_Archive-Requirements-Survey_20070619.pdf
5. Shearer, K. (2009) “Survey of digital preservation practices in Canada. final report”, Library and Archives of Canada available at: www.collectionscanada.gc.ca/digital-initiatives/012018-3100-e.html
6. The Information Management Journal (2002), “Preservation of the times”, 36(2, March/April, pp. 38-43, available at: www.panix.com/~jeffr/Prof/Pubs/DigitalLongevity/arma.paper.from-journal.pdf
7. Slats, J. “Digital preservation testbed white paper emulation: context and current status”, available at: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.104.9117> accessed 1 February, 2010).
8. Granger, S. “Emulation as a digital preservation strategy”. D-Lib Magazine, 6(10, October, available at: www.dlib.org/dlib/october00/granger/10granger.html

9. International Standardization Organization (ISO) (2001), "ISO 15489-1:2001 information and documentation – records management – part 1: general", available at: www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=31908
10. International Standardization Organization (ISO) (2003), "ISO 14721:2003 space data and information transfer systems – open archival information system - reference model", available at: www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=24683
11. Digital Curation Centre (DCC) (n.d.), "The DRAMBORA Project", available at: www.dcc.ac.uk/resources/tools-and-applications/drambora
12. InterPARES 2010. "International Research on Permanent Authentic Records in Electronic Systems", available at: www.interpares.org
13. Scielo. Available at: www.scielo.org
14. e-LAC. Available at: www.eclac.org/socinfo/elac/default.asp?idioma=IN
15. InterPARES. (2008) "International Research on Permanent Authentic Records in Electronic Systems. 'Team Brazil', 'Team Mexico', 'Team Colombia'", available at: www.interpares.org/ip3/ip3_index.cfm
16. UNESCO Archives Portal, available at: www.unesco-ci.org/cgi-bin/portals/archives/page.cgi?g=Resources%2FPreservation_and_Conservation
17. Kenney, A. and Buckley, E. 2005. "Developing digital preservation programs: the cornell survey of institutional readiness, 2003-2005". "RLG Diginews". RLG's Online Newsletter for Digital Imaging and Preservation, August 15, Chart 12, available at: www.rlg.org/en/page.php?Page_ID=20744
18. Professional Information & Security Communities. 2008. "Top information security risks for 2008", available at: www.iso27001security.com/Top_information_security_risks_for_2008.pdf
19. Berners-Lee, T. (2001). "The Semantic Web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities", "Scientific American" May 17.
20. "In the Semantic Web data itself becomes part of the web and is able to be processed independently of application, platform, or domain. This is in contrast to the worldwide web as we know it today, which contains virtually boundless information in the form of documents. We can use computers to search for these documents, but they still have to be read and interpreted by humans before any useful information can be extrapolated. Computers can present you with information but can't understand what the information is well enough to display the data that is most relevant in a given circumstance. The Semantic Web, on the other hand, is about having data as well as documents on the Web so that machines can process, transform, assemble, and even act on the data in useful ways". Taken from Altova Web site: Available at: www.altova.com/semantic_web.html
21. Digital Libraries Federation. DLF. (2002). "Benchmark for faithful digital reproductions of monographs and serials", available at: www.diglib.org/standards/bmarkfin.htm#reproduction

References

Beagrie, N. (2003), *National Digital Preservation Initiatives: An Overview of Developments in Australia, France, The Netherlands, and the United Kingdom and of Related International Activity: Commissioned for and sponsored by the National Digital Information Infrastructure and Preservation Program, Library of Congress, Council on Library and Information Resources (CLIR), Washington, DC, April*, available at: www.clir.org/pubs/reports/pub116/pub116.pdf

- Bohn, R. and Short, J. (2009), *How much information 2009: report on American consumers*, Global Information Industry Center, University of California at San Diego, San Diego, CA, December 9, available at: http://hmi.ucsd.edu/pdf/HMI_2009_ConsumerReport_Dec9_2009.pdf
- Digital Libraries Federation (2006), "Preservation initiatives", *The Global Digital Format Registry*, available at: www.diglib.org/preserve.htm#initiatives
- Duranti, L. (2005), *La Conservación a largo plazo de documentos electrónicos auténticos: Hallazgos del proyecto InterPares*, Concejalía de Cultura: 3000, Cartagena, (Tendencias; 4).
- InterPARES (2006), *International Research on Permanent Authentic Records in Electronic Systems. 'Terminology Database'*, available at: www.interpares.org/ip2/ip2_terminology_db.cfm
- Lessig, L. (2001), *The future of ideas*, available at: www.the-future-of-ideas.com/
- Lessig, L. (2004a), *Free Culture*, Penguin Press, New York, NY, available at: www.free-culture.cc/freeculture.pdf
- Lessig, L. (2004b), *Cultura Libre*, available at: www.elastico.net/archives/001222.html
- Lyman, P. and Varian, H. (2000), "How much information?", *Journal of Electronic Publishing*, Vol. 6 No. 2, December, available at: <http://quod.lib.umich.edu/cgi/t/text/text-idx?c=jep;view=text;rgn=main;idno=3336451.0006.204>
- Lyman, P. and Varian, H. (2003), *How much information 2003?*, University of California at Berkeley, Berkeley, CA, available at: www2.sims.berkeley.edu/research/projects/how-much-info-2003/
- Schneier, B. (2008), *Schneier on Security*, Wiley, New York, NY.
- Solove, D. (2007), *The Future of Reputation*, Yale University Press, New Haven, CT.
- Verzosa, F.A. (2005), "Digital initiatives in archival preservation", paper presented at the International Conference on Challenges in Preserving and Managing Cultural Heritage Resources, Institute of Social Order, October 19-21, Ateneo of Manila University, Quezon City.
- Zittrain, J. (2008), *The Future of the Internet and How to Stop It*, Yale University Press, New Haven, CT, available at: <http://futureoftheinternet.org/static/ZittrainTheFutureoftheInternet.pdf>

Further reading

- Duranti, L. (2005), *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project (San Miniato: Archilab, 2005). Authored the Acknowledgements and Introduction, and the Strategy Task Force Report (see Chapters/Parts of Books); co-authored the Authenticity Task Force Report.*
- Duranti, L. and Thibodeau, K. (2005), "The concept of record in interactive, experiential and dynamic environments: the view of InterPARES", *Archival Science*, Vol. 5 Nos 2-4, pp. 13-68.
- UNESCO. Memory of the World Program (2009), *Guidelines for the Nomination of Digital Documentary Heritage for Inclusion in the Memory of the World Registers*, UNESCO, Paris.
- Voutsas-M, J. (2009), *Preservación del Patrimonio Documental Digital en México*, UNAM: Centro Universitario de Investigaciones Bibliotecológicas, Mexico City, available at: http://132.248.242.3/~publica/archivos/libros/preservacion_patrimonio.pdf
- World Intellectual Property Organization (WIPO) (Organización Mundial de la Propiedad Intelectual) (n.d.), available at: www.wipo.int/portal/index.html.en and www.wipo.int/portal/index.html.es

About the author

Juan Voutssas, a PhD and Master's Degree in Library and Information Sciences, is an undergraduate Computer Systems Engineer. He has more than 32 years of experience in information and communications technologies (ICTs). His current academic job is as a Researcher at the "University Center for Library Sciences Research" (CUIB) at the National Autonomous University of Mexico (UNAM). Juan has been a Professor in undergraduate and graduate levels in the School for Library Sciences at UNAM for 25 years. He is current Director of the Mexican TEAM of the International International Research about Permanent Authentic Records in Electronic Systems (InterPARES) project. He is the former President of the National Academy of Librarians of Mexico and was a Member of the Academic Computing Advisory Council at UNAM for 18 years. He is a former Member of the Supercomputing Advisory Council at UNAM. His former managing jobs at UNAM were: Director of the Software Department at the Directorate for Academic Computing Services, Director of the Science and Humanities Information and Documentation Center, Director of the Computing Services Administrative Center, and Vice-director of Computing Services at the Directorate for Libraries. He has more than 50 works in books, papers, conferences, congresses, etc. Juan is a registered researcher – SNI – within the Governmental National Science and Technology Council (CONACyT). Juan Voutssas can be contacted at: voutssas@unam.mx