

International Journal of Advance Research in Computer Science and Management Studies

Research Article / Survey Paper / Case Study

Available online at: www.ijarcsms.com

Challenges in Preservation of Electronic Information

Arun Kumar H S¹Assistant Librarian
Surana College Pg Kengeri,
Bangalore-60, India**Lokesha M²**Guest Assistant Librarian
University Library
Mangalore -574199, India

Abstract: The difficulty and expense of preserving digital information is a potential impediment to digital library development. Preservation of traditional materials became more successful and systematic after libraries and archives integrated preservation into overall planning and resource allocation. Digital preservation is largely experimental and replete with the risks associated with untested methods. Digital preservation strategies are shaped by the needs and constraints of repositories with little consideration for the requirements of current and future users of digital scholarly resources. This article discusses the present state of digital preservation, articulates requirements of both users and custodians, and suggests research needs in storage media, migration, conversion, and overall management strategies. Additional research in these areas would help developers of digital libraries and other institutions with preservation responsibilities to integrate long-term preservation into program planning, administration, system architectures, and resource allocation.

Preservation of Electronic Information has always been a challenge. No matter due technology restriction or costs. Rapid growth and vary in technology have made creating electronic contents and publish a lot easier and preserving seems invisible but challenges are in advance of electronic information preservation. The aim of this paper is to identify the issues and project the future of electronic information preservation

Keywords: Electronic Information Preservation, Digital Technology, Cost of Electronic Information Preservation.

I. INTRODUCTION

When we deal preservation, the first and probably the only concept we have in mind is preservation the loss of materials but in fact it applies to easy, quick and cheaper access of the materials for current and future users. One of the challenges in digital preservation is that it is highly dependent on various software and hardware and as there is a very close connection between technology and digital preservation, there are unsolved disagreements experts on among experts on this. Since the advances in technology have become quicker than any time in history, the limitation in usage and transferring of digital materials into new formats is one of the challenges that digital preservation encounters with it. One of the suggested solutions is using simple formats which even modern software and hardware can retrieve them easily. An example is doc formats which is presented by Microsoft Office to save text files. Since a new format called “docx” was brought into life in Office Suit 2007, if a file is saved in “docx” format, retrieving this file by previous version of Microsoft Office 2003 would face difficulties especially if it includes tables or graphics.

Digital preservation is the series of management policies and activities necessary to ensure the ensure usability, authenticity discoverability, and accessibility of content over the very long term. While backup, system redundancy, and byte replication may be used by delivery organizations and digital preservation organizations, these actions alone are not sufficient for digital preservation.

II. DEFINITION OF TERMS

Key terms used throughout this paper are defined below. In some cases, these definitions are for consistency within the presentation and are not indicative of general consensus within the community.

Born digital – materials that are created in bits and bytes rather than being digitized from paper or other analog medium.

Digital archiving – storing the digital information (e.g., creating an institutional repository or digital archive)

Digital preservation – keeping the bits and bytes safe and unaltered for a long period of time

Digitization – converting materials in non-digital form (analog) such as paper, to digital form

Emulation – running old products by recreating the environment of the old hardware and software without actually using the old hardware and software.

Migration – moving a digital product from one version of a program, operating system or hardware environment to another over time.

Permanent or Long-term access – the ability to use a preserved object long after its initial preservati

Recapturing – copying the content from the original resource again in order to ensure that changes made to the resources are incorporated in the archival version

Refreshing – moving a digital object to a new instance of the same storage medium, retaining the same Operating system and hardware environment.

III. WHAT IS PRESERVATION PLANNING?

Libraries, archives, museums, and historical societies are responsible not only for collecting, interpreting, and exhibiting significant materials that document history, but also for the long-term preservation, security, and accessibility of these materials. The American Association of Museums recognizes this responsibility. It states in its Code of Ethics for Museums that a museum must insure that the "collections in its custody are protected, secure, unencumbered, cared for, and preserved." Preservation is an integral part of a cultural institution's mission, and preservation planning should be part of its overall strategic plan.

- » Preservation planning is a process by which the general and specific needs for the care of collections are determined, priorities are established, and resources for implementation are identified.
- » Its main purpose is to define a course of action that will allow an institution to set its present and future preservation agendas.
- » In addition, it identifies the actions an institution will take and those it probably will never take so that resources can be allocated appropriately.

IV. PROBLEMS IN DIGITAL PRESERVATION

The Uniform outlook will help the reader to follow the proceedings. This can be obtained most easily if authors use this template file to construct their papers. Please note the following details: this template is an A4 format with 20 mm margins left, right, top and bottom. Header and footer shall be positioned 16 mm from the edge. All text paragraphs should be single spaced, with first line intended by 10 mm. Double spacing should only be used before and after headings and subheading as shown in this example. Position and style of headings and subheadings should follow this example. No spaces should be placed between paragraphs.

The lack of adaption is one of the problems that Preservation of Electronic Information has to encounter. For example if a file is created in Windows 95, there is a probability that retrieving and reading of that file in Windows 7 will face difficulties. So, the preservation itself is a lot easier but the process of retrieving is not that accessible and the reason is the extreme rapid changes and developments in technology. According to a survey by Spinellis which is done on reference URLs of articles in Computer and Communications journal, he found out that 28% of reference links were not accessible by year 2000 and this soared to 41% by 2002, such issues have raised concerns in digital preservation.

Although repository media like “CD”s have high capabilities in storing big amount of data comparing to paper or microfilm but on the contrary, their “life” is too short comparing to paper and microfilm. The maximum age for optical disks is 25 years while it is 500 years for microfilm. According to RLG report on reliable digital repositories, they should have two main features: The repository should involve with proper standards and technologies that prepare the structure for digital preservation. The repository is a system containing software and hardware that can be certain about following specific rules.

V. MAIN METHODS IN PRESERVATION OF ELECTRONIC INFORMATION

Libraries and archiving centres use three main approaches in electronic information

Preservation. 1) Refreshing 2) Migration 3) Emulation

1) Refreshing

Refreshing is the transfer of data between two types of the same storage medium so there are no bitrot changes or alteration of data. For example, transferring census data from an old reservation CD to a new one. This strategy may need to be combined with migration when the software or hardware required to read the data is no longer available or is unable to understand the format of the data. Refreshing will likely always be necessary due to the deterioration of physical media.

2) Migration: Migration is the transferring of data to newer system environments (Garrett et al., 1996). This may include conversion of resources from one file format to another (e.g., conversion of Microsoft Word to PDF or Open Document) or from one operating system to another (e.g., Windows to GNU/Linux) so the resource remains fully accessible and functional. Two significant problems face migration as a plausible method of digital preservation in the long terms. Due to the fact that digital objects are subject to a state of near continuous change, migration may cause problems in relation to authenticity and migration has proven to be time-consuming and expensive for "large collections of heterogeneous objects, which would need constant monitoring and intervention. Migration can be a very useful strategy for preserving data stored on external storage media (e.g. CDs, USB flash drives, and 3.5" floppy disks). These types of devices are generally not recommended for long-term use, and the data can become inaccessible due to media and hardware obsolescence or degradation.

3) Emulation: In this method, the format of the file is not converted but an environment is designed to run the files in exactly the same system they were run in the past. A clear example is designing a software application which can run Microsoft Windows software in Linux. The main aim of this simulation is keeping the existing file as is and if the process is performed properly, it can be said that the goal is achieved. Not only the exact retrieving of the files are considered, but also the “feeling” which means retrieving the appearance of the file like frame, color, etc.

Another example of emulation as a form of digital preservation can be seen in the example of Emory University and the Salman Rushdie's papers. Rushdie donated an outdated computer to the Emory University library, which was so old that the library was unable to extract papers from the hard drive. In order to procure the papers, the library emulated the old software system and was able to take the papers off his old computer.

VI. COMPARISON OF MIGRATION AND EMULATION METHODS

One of the limitations of Emulation method is that because of technical limitations, it is not possible to make Simulator for every program. On the other hand, Migration costs huge amount of financial sources and needs life-time support because this migration has to continue as long as digital technologies keep changing.

VII. MAIN FEATURES OF PRESERVATION OF ELECTRONIC INFORMATION

In association with digital preservation it is important to take in account that data must be saved with no change or loss. Access, running and understanding must be achieved by the user which means searching and retrieving the information must be provided. The above mentioned goals must be available in long-term. Alongside the all efforts put in to providing these aims, digital preservation is facing enormous challenges. Reaching above goals needs skilled people in preservation and since they should be skillful in traditional preservation and computer science, it will take lots of resources. According to a survey, the lack of clarity in responsibilities is the major barrier in digital preservation. A suggested solution to overcome this problem is using organizations and institutes which are experts in preservation. This solution's limitation is that because of limited financial sources they have, there is not great assurance on them for long-term support and as a result. Long-term relationship cannot be established.

VIII. TYPES FOR PRESERVATION OF ELECTRONIC INFORMATION

Speaking of digital preservation, the immediate concept is converting books, magazines, reports etc., which are in paper into digital format but there is another type of digital preservation which is preserving "Digital Objects". For example: Web Pages, digital images, etc. some arguments are running on this topic, such as how we should preserve a document or a report which is in paper form? Should we type the contents or scan the pages? Should we care about the contents or the appearance? Opponents of scanning method argue that what if some parts of the document were unclear to read? Should we create a file containing the Contents of the document alongside scanning the pages? What about the costs?

- » Cost for Converting Information into Digital Form
- » Cost Variables in the Digitization Process
- » Costing Studies and Comparative Analysis
- » Costs for Maintaining Digital Information

IX. CONCLUSION

Although with the rapid development of technologies especially in digital world new opportunities are made for easier preservation of physical and digital materials but there are some issues which remain unsolved about costs and stability of digital technologies. All these questions must be taken to account when planning a digital preservation project. Proponents of reducing costs in digital preservation should face this question that how much limitation should be applied to costs when it's the matter of preserving cultural heritage? Is it possible to rate a nation's cultural heritage? If the answer to the above question is negative, then is it imaginable to avoid digital preservation? The Great Fire of London in 1660 destroyed lots of books and documents that only one volume existed and they are gone for good. By the rapid developments in technology, access to recording tools like camera-equipped cell phones, more and more documents are being produced which are not preserved in libraries or achieving centre's. For example; photos taken by ordinary people during natural disasters such as Hurrigan Katrina in 2005 or Tsunami in Indonesia in 2004. Probably the best way to preserve these photos is digital preserving.

References

1. Rothenberg, J. (2000). "Preserving Authentic Digital Information", in Authenticity in a Digital Environment, pp. 51-68
2. Meijer, A. (2001). "Electronic Records Management and Public Accountability: Beyond an Instrumental Approach", in The Information Society, no 17, pp. 256-270.
3. Kalbitzer, S. (2004). Novel Concepts for Mass Storage of Archival Data. Nuclear Instruments and
4. Methods in Physics Research B 218, pp 343-354
5. Marcum, D. B. (1997). "A Moral and Legal Obligation: Preservation in the Digital Age", in International Information and Librarian Review 29 , pp. 357-365.
6. Rezgui, A., Bouguetta, A. & Eltiweissy, M. Y. (2003). Privacy on the Web: Facts, challenges, and solutions. IEEE Security and Privacy , November/December, 40-48.
7. Marcum, D. B. (1997). A moral and legal obligation: Preservation in the digital age. International Information and Librarian Review, 29 , 357-365.
8. Asproth, V. (2005). Information technology challenges for the long-term preservation of Electronic information . International Journal for Public Information Systems.
9. Digital Preservation Coalition. (2002). [Online]. Available: <http://www.dpconline.org/graphics/index.htm> [21 April 2004].
10. Digital Preservation Coalition. (2004) DPC/PADI What's new in digital preservation. [Online]. Available: <http://www.dpconline.org/graphics/whatsnew/> [21 June 2004].

AUTHOR(S) PROFILE

ARUN KUMAR H S MLISc (Master of Library and Information Science) ASST. LIBRARIAN,
SURANA COLLEGE PG Departments KENGERI, BANGALORE-60, KARNATAKA, INDIA.



LOKESHA M MLISc (Master of Library and Information Science) ASST. LIBRARIAN
(Guest) Mangalore University Library Mangalore- 574199, KARNATAKA, INDIA.