



Digital Preservation of Information Resources: Issues and Strategies

DR. RAJIV VIJ
Assistant Librarian,
Ch. Devi Lal University, Sirsa
Haryana (India)

Abstract:

Digital preservation combines policies, strategies and actions that ensure access to digital content over time. An information resource refer to a system resource in any component of limited availability within a computer system, a web resource, a data source accessible at the World Wide Web and an electronic resource, information which can be stored in the form of electrical signals. The purpose of preservation of information sources is to protect information of enduring value for access by present and future generations. In recent decades, many major libraries and information centres and archives established formal preservation programs for traditional materials in paper, microform, photographic, and to lesser degree audio-visual formats. Digital preservation is largely experimental and replete with the risks associated with untested methods. Digital preservation adds a new set of challenges for libraries and archives to the existing task of preserving a legacy of materials in traditional formats. But the strategies, methods, and technologies for long-term preservation of digital information that do exist are not necessarily feasible technologically for preservation on a mass scale, nor are they affordable given the vast quantities of digital information being generated.

This paper discusses the types of information resources, present state of digital preservation, articulates requirements of both users and custodians, and suggests with limitations and recommendation for research needs in storage media, migration, conversion, and overall management strategies.

Keywords: *Digital preservation, Digital content, Information resources, Strategies*

1. Introduction

Digital preservation combines policies, strategies and actions that ensure access to digital content over time. At present there is as yet no viable long-term strategy to ensure that digital information will be readable in the future. Not only are digital documents vulnerable to loss via media decay and obsolescence, but they become equally inaccessible and unreadable if the software needed to interpret them, or the hardware on which that software runs, is lost or becomes obsolete. The purpose of preservation of information sources is to protect information of enduring value for access by present and future generations. Efforts are made to discuss the concept of digital preservation, information resources and its types, present state of digital preservation, articulate requirements of both users and custodians and suggest recommendations for research needs in storage media, migration, conversion and overall management strategies for digital preservation of information resources.

2. Definitions

2.1 Digital Preservation

Digital preservation is a process by which data is preserved in digital form in order to ensure usability, durability and intellectual integrity of the information contained therein.

Kelly (1999), defines digital preservation as “the storage, maintenance, and accessibility of digital object (include any digital material such as a text document, an image file, a multimedia CD-ROM or a database) over long term, usually as a consequence of applying one or more digital preservation strategies. The term digital preservation is used in different ways.

Fresco (1999) defines the term as, “the storage, maintenance and access to digital objects over long term”. The key point of this definition is that it is about ensuring that intellectual content, which is already in digital form, remains accessible to the future generations.

As Marcum noted in 1996, "Preservation is a fundamental responsibility of libraries and archives of record. To be sure, the preservation imperative has been imperfectly carried out in the print environment, but the problem grows even more complicated in the digital world."

According to ALA (2007), “Digital preservation combines policies, strategies and actions that ensures access to digital content over time”.

2.2 Information Resources

As a dictionary meaning “Information resources are defined as the data and information used by an organization”.

According to wikipedia information **resource** may refer to:

-A system resource in computer science, any component of limited availability within a computer system.

-A web resource, a data source accessible at the World Wide Web.

-An electronic resource, information which can be stored in the form of electrical signals

According to Encyclopaedia of Networked and Virtual Organizations, “Information resource is an element of infrastructure that enables the transaction of certain selected significant and relevant data, prepared so as to provide content and information services that can be used directly by the user”.

According to Encyclopaedia of Information Science and Technology, Information resources means, “A collection of valuable information generated by human activities. In a broader sense, it also includes related equipment, personnel, and capital.”

3. Types of Information resources

3.1 Published Work

Published works include Government publications like;

Journals - A journal is a periodical, which generally contains materials relating to research. It appears at regular intervals-weekly, monthly or quarterly. Its content varies and can include editorials, articles, boo reviews etc.

Newspapers- As sources of current information and topical comment, newspapers can be good sources for primary research information. Newspapers are issued either daily, weekly or monthly. It contain news, opinions, advertisements and other subjects related to current affairs

Monographs and textbooks are Publications that deal comprehensively with a specific subject;

Reference works-These are Dictionaries, Encyclopaedias, Biographies, Yearbooks and address books etc.

Audio Visual-Other media such as audio cassettes or videos

Electronic media- Information that is electronically available - GD Rom programs, e-Books, e-Journals etc.

3.2 Un-Published Work

Un-published **works** includes;

Human resources - Lecturers, Colleagues and others

Dissertations/Theses- These are major sources of primary research output. Some of the most current and original research every year is produced by postgraduate researchers as part of an academic course for a higher degree and copy usually made available in library of an University

Reports – Reports are produced by agencies and departments on specific topics or issues.. These are written description of a completed research project or an interim progress report

Grey Literature- Information that is not available through the normal book selling channels such as reports, manuscripts and patents

Standards

Standards are consensus agreements drawn up by representative collections of people who have a particular interest in the subject. These might be manufacturers, users, research organisations, or government departments.

4. Background

Preservation of information resources has always been and will continue to be a source of concern to libraries, except a lasting solution is found. The philosophy and practice of preservation has roots in many earlier traditions. In general preservation, as a formal profession in libraries and archives dates from the twentieth century. In 1933, William Barrow introduces the field of conservation to paper de-acidification when he publishes a paper on the acid paper problem. In 1966, The flood of the River Arno in Florence, Italy damaged or destroyed millions of rare books and led to the development of restoration laboratories and new methods in conservation. Paul N. Banks (1934–2000) published regularly on bookbinding, book and paper conservation, and problems related to conservation. Ellen R. (1983) edited *The Brown Book: A Directory of Preservation Information*. McCrady 2002-2008 – edited and published the "Abbey Newsletter", covering important information for preservation professionals. Paul Conway 2005 – Conway is an associate professor in the University Of Michigan School Of Information. His research and educational work focuses primarily on digital preservation and electronic media. Walter Henry 2007 – Henry, a conservator at the Stanford University Libraries and Academic Information Resources, is the creator of Conservation On Line and the Conservation DistList. Janet Gertz 2008 – Director for Preservation, Columbia University Libraries has been chair of ALCTS' Preservation and Reformatting Section (PARS).

5. Objectives

The purpose of preservation is to protect information of enduring value for access by present and future generations Conway, (1990). In according to an institution's achievable resources, the main objectives of a digital preservation policy are to achieve and preserve any type of digital record that is:

- created using any type of application;
- created on any computing platform;
- delivered on any digital media;

- preserving and providing continued access to digital material, both born digital and digitised material for current and future generations;
- Continue to develop expertise in the creation and management of digital collections and Information
- Adapt preservation strategies to incorporate the capabilities afforded by new and emerging technologies in cost-effective and responsible way
- ensuring that preserved digital materials are authentic;
- preserving damage and deterioration of the physical media by ensuring an environmental control;
- Contribute to the development of the standards and practice of the digital preservation community

According to Abby Smith (1998), “The goal of any preservation program is to ensure long- term, ready access to the information resources of an institution”.

6. Issues and Challenges

Digital preservation is not a single task but encompasses all strategies and measures taken to access, maintain and save digital documents to avoid their loss. These measures besides raising non technical issues concerning jurisdiction, funding, responsibility for successive phases of the digital document life cycle, and the development of policies requiring adherence to standard techniques and practices also raise technical, social and legal issues adding to the difficulty of the task to prevent the loss of digital information. These include:

6.1 Technological Obsolescence

In addition to the technical issues, many of the issues surrounding digital preservation involve an institution’s willingness to manage their digitized data Teper, (2005). It must be acknowledged that digital technology is not the only answer, and that hard copies of artefacts and materials are still very important. Acid-free paper and microfilm can last for 500 years in a library that is air conditioned and dry, but digital files become obsolete every few years at best (Brand, 1999). Quickly changing digital technologies are one of the most serious challenges associated with digital preservation. Because there is no way of knowing what formats or procedures people will be using 50 or 100 years from now, it is more important to focus on a shorter time frame. If professionals focus on the coming five, ten or even twenty years, they are more likely to have a better idea of what the practices will be (Gladney 2009). Software corporations are both part of the problem and the solution, as due to commercial interest, the devices, software and formats used to store information are often designed for obsolescence. In other words, corporations that design these methods design them with a shelf life in order to ensure the sales of latest upgrades and models (Reyes, 2013).

6.2 Migration of Digital Information

Planning for migration is difficult because there is limited experience with the types of migrations needed to maintain access to complex digital objects over extended periods of time. When a custodian assumes responsibility for preserving a digital object it may be difficult to predict when migration will be necessary, how much reformatting will be needed, and how much migration will cost. There are no reliable or comprehensive data on costs associated with migrations, either for specific technologies and formats or for particular collections, and little research underway on methodologies that would reduce the costs and burdens of migration.

6.3 Legal and Organizational Issues

Copying might occur as part of one of the digital preservation strategies, or it could merely be the process of copying digital information from a storage medium into the RAM of a computer. The exclusive right to copy an item, however, is one of the rights given to an author by copyright law. The author has other exclusive rights, including the rights to:

- prepare derivative works based upon the copyrighted work
- distribute copies of the copyrighted work to the public

- perform some copyrighted works publicly
- display some copyrighted works publicly
- control access to a work protected by the use of a technological measure

Digital preservation strategies may impinge on these rights. Migration, for example, may be a violation of the copyright owner's right to prepare a derivative work. Making a digital work broadly available may impinge on the copyright owner's distribution, performance, and display rights. It is imperative, therefore, that digital preservation programs remain in close contact with their institution's legal advisors to ensure that they do not place their institution at an unacceptable level of risk.

6.4 Infrastructure

For digital preservation, the organizational effort is the process of building infrastructure-necessarily involves multiple, interrelated factors such as, institutions, services, technologies, and qualified personnel capable of supporting a distributed system of digital archive. The effort to meet cultural imperative of digital preservation requires a complex interaction and re-interaction.

6.5 Conceptual Framework or Standards

Another issue is the absence of established standards, protocols, and proven methods for preserving digital information. There is lack of adequate research in the areas such as framing key problems associated with digital information, defining critical issues, establishing standards for digital preservation, etc. Hedstrom & Montgomery (1998), in their survey 'Digital Preservation Needs and Requirements in RLG Member Institutions' found that majority of member institutions have not developed policies for the acquisition, storage, refreshing and migration of digital information and have not developed methods to preserve digital information.

In recent decades, many major libraries and archives established formal preservation programs for traditional materials in paper, microform, photographic, and to lesser degree audio-visual formats. Preservationists within the library and archival community have been instrumental in developing an array of tools and methodologies to reduce the decay of traditional materials and to restore books and documents that have deteriorated to such an extent that their longevity and usability are threatened. These programs represent a significant risks and challenge for both government agencies and archival institutions. Factors affecting the level of risk include:

- The carriers used to store information resources are usually unstable and deteriorate within a few years or decades at most.
- The use of digital records and other information resources requires specific combinations of hardware and software that typically become obsolete after a few years, rendering the digital records inaccessible.
- File formats change over time, which can result in inaccessibility in using current software.
- File formats are sometimes unable to be determined, especially for older software.
- Possible risks of loss of Digital information resources in the event of disasters such as fire, flood, equipment failure, or virus or direct attack that disables stored data and operating systems.
- Access barriers such as password protection, encryption and security devices may prevent ongoing access beyond the circumstances for which they were designed.
- The digital records may be well protected, but so poorly identified and described that potential users cannot find them.
- So much contextual information may be lost that the records themselves are unintelligible or not trusted.

The inherent instability and vulnerability of digital records affects the ways in which the preserving institutions and archives secures, manages and preserves digital records and other information resources. Digital preservation also adds a new set of challenges for libraries and archives to the existing task of preserving a legacy of materials in traditional formats. Preservation programs include

administrative and technical components, such as hiring staff with expertise in preservation administration, using preventive measures to arrest deterioration of materials, taking remedial actions to restore the usability of selected materials, and incorporating preservation needs and requirements into overall program planning and resource allocation. More insidious and challenging than media deterioration is the problem of obsolescence in retrieval and playback technologies. Innovation in the computer hardware, storage, and software industries continues at a rapid pace, usually yielding greater storage and processing capacities at lower cost. Digital works which are created using new or emerging software applications are especially vulnerable to software obsolescence because standards for encoding, representation, retrieval, and other functions take time to develop.

6.6 Digital Preservation Standards

Standards facilitate the access, discovery and sharing of digital resources, as well as their long-term preservation. There are both generic standards applicable to all sectors that can support digital preservation, and industry-specific standards that may need to be adhered to. . Long-term access to any type of asset, regardless of format, depends upon institutional commitment and collection care. Sustained access to digital material requires that the preserving institution has the infrastructure, capacity, skills and resources to enable it to keep abreast with the fast moving digital environment. Using standards that are relevant to the digital institutional environment helps with organisational compliance and interoperability between diverse systems within and beyond the sector. But digital preservation is constrained by the absence of established standards, protocols, and proven methods for preserving digital information and by the tendency to consider preservation issues only at the end of a project or after a sensational loss. Levy and Marshall (1995) observed that "With few exceptions, digital library research has focused on architectures and systems for information organization, retrieval, presentation, and visualization, and on the administration of intellectual property rights"

A standards based approach to digital preservation is important, but there are also factors which inhibit their use as a digital preservation strategy. The pace of change is so rapid that standards which have reached the stage of being formally endorsed - a process which usually takes years - will inevitably lag behind developments and may even be superseded. Competitive pressures between suppliers encourage the development of proprietary extensions to, or implementations of standards which can dilute the advantages of consistency and interoperability for preservation. The standards themselves adapt and change to new technological environments, leading to a number of variations of the original standard which may or may not be interoperable in the long-term even if they are backwards compatible in the short-term. Standards can be intimidating to read and resource intensive to implement. In such a changeable and highly distributed environment, it is impossible to be completely prescriptive. These factors mean that standards will need to be seen as part of a suite of preservation strategies rather than the key strategy itself. The digital environment is not inclined to be constrained by rigid rules and a digital preservation programme can often be a blend of standards and best practice that is sufficiently flexible and adapted to suit the needs of the organisation, its circumstances and the digital materials being managed.

6.7 Digital Preservation Strategies

The aim of the strategy is to build upon the existing digital preservation work to continue to enable the long-term sustainability of digital assets, both born digital and digitised, held by the Library.

Lyman and Besser noted, "The long term preservation of information in digital form requires not only technical solutions and new organizational strategies, but also the building of a new culture that values and supports the survival of bits over time."

Beagrie, Semple, Williams, and Wright reinforced the idea that "...any long-term access and future benefit may be heavily dependent on digital preservation strategies being in place and underpinned by relevant policy and procedures...and that digital preservation policy should be integrated into

business drivers, activities and functions e.g. regulatory compliance, staff development, applied technology, academic excellence."

At the moment, primarily migration and emulation may be identified as relevant digital preservation strategies. Refreshing has to be rated as a supplementing strategy. The approach to set up a "hardware museum" seems to be hardly suitable to provide a long-term perspective or strategy for digital preservation.

6.8 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 preservation 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. There are certain advantages of refreshing strategy which includes:

- The risk of "data carrier failure" is dealt with.
- The risk of "availability of reproducers" is dealt with.
- Relatively inexpensive to implement
- Relatively easy to implement"

But there are certain problems associated with the strategy are:

- Problem of dependence on software not solved
- Problem of format ageing not solved
- Not sustainable without combination with other digital preservation strategies

6.9 Migration

Migration is the transferring of data to newer system environments 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 Linux) so the resource remains fully accessible and functional. It is necessary to Store formats in an up-to-date format before they become obsolete. The associated advantages are:

- The risks of "availability of operating systems" and "availability of software" are dealt with.
- The risk of format ageing is dealt with.
- The information is preserved in a new, contemporary format.

But there are possible disadvantages associated with this strategy are:

- It is Very expensive due to assessment of formats and observation of the development of formats
- Continuous task which needs to be repeated on a regular basis
- As a stand-alone procedure, it does not protect against obsolescence of devices and data carriers.
- Information may be lost during migration

6.10 Emulation

Emulation is the replicating of functionality of an obsolete system. According to Van Der Hoeven, "Emulation does not focus on the digital object, but on the hard- and software environment in which the object is rendered. It aims at (re)creating the environment in which the digital object was originally created.". Examples are having the ability to replicate or imitate another operating system. The approach associated with this strategy is "Imitate" old hardware and software environment by means of an emulator. The advantages of this strategy are:

- The risks of "availability of operating systems" and "availability of software" are dealt with.
- The risk of format ageing is dealt with.
- Data can to a great extent be used as originally.

- Some of the drawbacks of this strategy includes;
- Very costly and expensive
- A large number of different emulators are required (depending on formats and required software).
- It is not certain that emulators will run on future operating and file systems.

According to Lesk (1992), Most librarians and archivists have accepted the basic wisdom that digital preservation depends upon copying, not on the survival of the physical media". But copying, also referred to as "refreshing" or "migration," is more complex than simply transferring a stream of bits from old to new media or from one generation of systems to the next. Complex and expensive transformations of digital objects often are necessary to preserve digital materials so that they remain legitimate representations of the original versions and useful sources for analysis and research (Preserving Digital Information). Current methods for preserving digital materials do not fully support achieving these objectives. When faced with the responsibility for preserving digital materials, archives and libraries face a series of complex and difficult choices based on the format of the original materials, the anticipated uses for it, and the technical and financial resources available to invest in preservation initiatives which range from very elementary and established methods to proposals that have not yet been tested.

Another strategy for digital preservation is to preserve digital information in the simplest possible digital formats in order to minimize the requirements for specific retrieval software and to avoid problems of software obsolescence. Digital information can be transferred across successive generations of technology in a "software-independent" format as ASCII text files or as flat files with simple, uniform structures.

7. Conclusion

The challenges in digital preservation involve dealing with not just the technologies of the past, but also those to come. The issues raised in digital preservation of information resources are of immense importance but the strategies that have been suggested including the standards and the migration of digital material into new forms are inadequate. Preservation environment buy time for careful decision making and represent highly tangible commitments to long term preservation. Collaboration is essential in the development of preservation repositories which should not depend for their existence on resources allocated by any single organisation. The digital preservation priorities should re-assert leadership regarding standards and best practices for digital collection building on par with the leadership that now exists in the digital preservation arena. As De Vorsey says, "the best that the preservation community can do with digital material is to make educated guesses based on a few decades of mostly anecdotal experience"

References

1. ALA, (2007). Definitions of digital preservation, Chicago, pres 0408.pdf American Library Association. Retrieved from: <http://www.ala/mgrps/divs/alcts/resources/preserve/defdig>
2. archival preservation. Library Resources & Technical Services, 49(1), 32-
3. Conway, Paul. (1990). "Digitizing Preservation." Library Journal, (February 1, 1994), 42-45.
4. De Vo De Vorsey, Kevin, Peter McKinney, 2010, "Digital preservation in capable hands: taking control of risk assessment at the National Library of New Zealand". Information Standards Quarterly, vol. 22, no. 2, pp. 41-44, 2010.
5. Deanna, B. Marcum, "The Preservation of Digital Information," The Journal of Academic Librarianship 22, No. 6 (November 1996): 452.
6. digital objects. The American Archivist, 72(2), 401-435.
7. Ellen, R. The Brown book: a directory of preservation information/National Trust for Historic Preservation in the United States :Published-Washington, D.C. : Preservation Press, c1983.
8. Encyclopaedia of Information Science and Technology, Second Edition,

9. Fresco, M.(1999). Digital preservation: state of the art. Conference Proceedings of DG Information Society, Luxembourg, 17th-19th November 1999
10. Gladney, H. M. (2009). Long-term preservation of digital records: Trustworthy
11. Hedstrom, Margaret & Montgomery, (1991) Shen. Digital preservation needs and requirements in RLG member libraries: a study conducted by Research Libraries Group, December 1998. (Available at: <http://www.rlg.org/preserv/digpres.html>)
12. Lesk, Michael (1992). Preservation of New Technology: A Report of the Technology Assessment Advisory Committee to the Commission on Preservation and Access. Washington, D.C., Commission on Preservation and Access, 1992.
13. Levy, David M. and Catherine C. Marshall. "Going Digital: A Look at Assumptions Underlying Digital Libraries." Communications of the ACM, 58(4) (1995), 77-84
14. Lyman and Besser, "Defining the Problem of Our Vanishing Memory," 12.
15. McCrady, Mold: The Whole Picture - Effect of Mold on Schools, Homes, & Human Beings, Part 4." Abbey Newsletter vol. 23, no. 7-8 (1999), <http://cool.conservation-us.org/byorg/abbey/an/an23/an23-7/an23-702.html>
16. Neil Beagrie, Najla Semple, Peter Williams, and Richard Wright, Digital Preservation Policies Study Part 1: Final Report October 2008, A Study Funded by JISC (Higher Education Funding Council for England (HEFCE): October 2008)
17. Reyes, V. (2013). We created it, now how do we save it? Issues in preserving personal information, a review. Preservation, Digital Technology & Culture, 42(3), 150-154. doi:10.1515/pdte-2013
18. Russel, Kelly. Digital preservation: ensuring access to digital materials into the future.1999 (Available at: www.leeds.ac.uk/cedars/chapter.htm)
19. Smith, Abby. (1998). "Preservation in the Future Tense." CLIR Issues 3 (May/June). Washington, D.C.: Council on Library & Information Resources.
20. Teper, T. (2005). Current and emerging challenges for the future of library and
21. Van der Hoeven "Emulation for Digital Preservation in Practice: The Results - van der Hoeven - International Journal of Digital Curation". ijdc.net.
22. Waters, Donald, and John Garret. (1996). Preserving Digital Information. Report of the Task Force on Archiving of Digital Information. Commissioned by the Commission on Preservation and Access and the Research Libraries Group, Inc. (May 1).