PRESERVATION OF DIGITAL INFORMATION IN LIBRARIES: ISSUES AND STRATEGIES

By

Bulu Maharana*
Dr. K. C. Panda**

<u>ABSTRCT</u>

The introduction of digital technology into the process of production, distribution and storage of information has made the libraries to go digital. Libraries all around the globe have responded to various challenges posed by the preservation of digital information and have encountered the technical, organizational, resource and legal issues associated with it. Although, the libraries have been experimenting with various preservation strategies such as technology preservation, emulation, migration etc., the need for a technologically feasible, financially affordable and widely acceptable strategy for preservation is still there. The paper discusses various issues and challenges associated with digital preservation and examines different strategies of digital preservation. The paper highlights initiatives undertaken at international level to develop workable approaches and best-practice preservation strategies for digital resources in libraries.

* Indian Institute of Management, Rajendra Nagar, Indore- 452012, Email: bulu@iimidr.ac.in
** Senior Lecturer, Deptt. Of Library and Information Science, Sambalpur University, Jyoti Vihar-768019, Sambalpur

0.   Introduction

Over recent years, libraries have grown increasingly reliant on digital materials. Libraries as information service providers have come to rely increasingly on digital information both as supplements to and parallels of print materials. Libraries are also encountering new resources, which are "born digital" and have no print or analogue equivalent- they exist only in digital form. The introduction of digital technologies into the processes of production, distribution and storage of information challenges the capacity and abilities of libraries, archives, and other cultural institutions to carry out their responsibility for preservation. The purpose of preservation is to ensure protection of information of enduring value for access by present and future generations (Conway, 1990; 206). Libraries, that traditionally have assumed the responsibility for preserving information, face technical, organizational, resource, and legal challenges in responding to the new demands for digital preservation.

1.   What is 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. A more precise definition as given by Kelly (1999), " 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.

1.1   Issues Relating to Digital Preservation

The fundamental problem of preserving electronic documents or "digital objects" stems from the nature of the objects themselves. Digital objects are accessible only through combinations of hardware and software. The hardware and software become obsolete in cycles of less than three years. Ensuring ongoing access, therefore, requires currency with technology changes, and moving digital objects from obsolete to current file formats, storage media, operating systems and so on.

A number of technical, social and legal issues add to the difficulty of the task. These include:

1.1.1 Technological Obsolescence

Innovation in computer hardware, storage and software industries continues at a rapid pace, usually yielding greater storage and processing capacities at lower cost. Devices, processes and software for recording and storing information are being replaced with new products and methods at regular three to five year cycles, driven primarily by market forces. Records created in digital form in the first instance and those converted retrospectively from paper or microfilm to digital form are equally vulnerable to technological obsolescence. It is costly and difficult for vendors to assure that their products are either *backwardly compatible* with previous version or that they can *inter-operate* with competing products.

1.1.2. Migration of Digital Information

Migration is the periodic transfer of digital materials from one hardware or software configuration to another or from one generation of computer technology to a subsequent generation. The purpose of migration is to preserve the integrity of digital objects and to retain the ability for clients to retrieve, display, and otherwise use them in face of constantly changing technology. It has always been a problem with the preservationists to make an exact copy or replica of a database or other digital object as

hardware and software change makes the object incompatible with the new generation of technology.

1.1.3 Legal and Organizational Issues

The barriers to decisive preservation action is caused by wide spread uncertainty about legal and organizational requirements for managing the intellectual property that digital information represents. Any preservation strategy will require the permission from the copyright holders. Permission could be needed to digitize a work of art; store the digitized image in an archive; move the digital image between archives; grant permission for authorization to the digital image; change file formats; change storage media; and also for the possible deletion of the digital image.

1.1.4 Infrastructure

For digital preservation, the organizational effort-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 iteration and reiteration.

1.1.5 Conceptual Framework or Standards

Another challenge 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, freshing and migration of digital information and have not developed methods to preserve digital information.

2.   Strategies for Digital Preservation

Digital preservation is defined as the storage, maintenance, and access to digital object over the long term, usually as a consequence of applying one or more digital preservation strategies. A digital preservation strategy should be evaluated against a set of criteria such as technological feasibility, cost-effectiveness, effectiveness in retaining the essential attributes of digital information, acceptance by creators, managers of digital repositories, and user communities (Hedstrom, 1999). The lack of technologically feasible and affordable method for digital preservation is a major obstacle for digital libraries.

Hendly (1998) identified three potential strategies for ensuring long term access to digital information. These include (1) Technology preservation, (2) Technology emulation and (3) Digital information migration. However, Russel (1999) and Bullock (1999) have accepted "output to permanent paper or microfilm" as a low-tech strategy for preservation.

2.1   Output to Permanent Paper

This particular strategy is sometimes referred as "change media" and there are those who advocate this as best long-term solution for the digital materials. As the name suggests, this approach involves printing out digital material and preserving the paper copy. In the short term, costs of producing the hard copies are also more predictable because the process can easily integrate into activities and structures for preserving traditional materials. However, this strategy, in long term, is unsuitable due to space constraints and availability of increasingly cost effective methods of digital storage. Reliance on hard copy is expensive, impractical and ultimately undesirable (Russel, 1999). Many digital objects simply can not be printed and would lose most of their unique functionality and attributes by printing. Complex multimedia CD-ROMs for example can simply not be printed with out losing many of the key characteristics of resources. Rothenberg (1999) rightly points out that 'printing a hard copy' is a simple but not a complete solution since some of the multimedia documents can not be printed at all and any interactivity, a document possess will be lost.

2.2   Technology Preservation/Preserve Technology

Another method for ensuring ongoing access to digital objects would be to simply keep older technology available for use. The most obvious way of ensuring that the object is preserved as it was created is to preserve the environment used to create and use resources, that is preserve the software and hardware environment that was used to access the resource when it was created. For some digital objects this may be the best solution-at least in the short-run because it ensures that the material is accessible by preserving the access tools as well as the object itself. However, in longer term this is more problematic. For example, issues of space and maintenance of hardware as well as the costs may make this an impractical solution in the long run. This method is used as an interim method when migration is not possible.

2.3   Technology Emulation

Emulation refers to creating new software that mimics the operations of older hardware or software in order to reproduce its performance. Thus not only are physical presence and content preserved, but digital objects could display original features (i.e. layout) and functionality available with the older software. Emulation has recently attracted attention as a potential strategy to assist preservation, recognizing that electronic material that is highly dependent on particular hardware and software will not lend itself to migration. Emulation is used to provide "backward compatibility" for digital objects in subsequent technological environments. Russel (1999) considers emulation as more like the 'just in time' option whereas, in technology preservation we will have the necessary hardware and software 'just in case'.

2.4   Digital Technology Migration

Migration is the primary strategy articulated by most organizations that plan to preserve digital objects. The Task Force on Archiving of Digital Information defines digital information migration as, "a periodic transfer of digital materials from one hardware or software configuration or from one generation of computer technology to a subsequent

generation." It covers a range of activities to periodically copy, convert or transfer digital information from one generation of technology to the subsequent ones. Migration may involve copying digital information from a medium that is becoming obsolete or physically deteriorating to a newer one (e.g., floppy disk to CD-ROM) and/or converting from one format to another (e.g., Microsoft Word to ASCII), and /or moving documents from one platform to another (MS DOS to UNIX) (Bullock, 1999).

Migration certainly preserves the physical presence and the content of a digital object. However, it may not preserve presentation, functionality and context. Successive migrations may eventually result in unacceptable data loss. Migration is undeniably an important strategy for preserving digital objects. However, it has yet to be tested and proven as a mechanism for managing complex multimedia objects in the long run.

3.   Global Projects on Digital Preservation

There are a number of good projects underway attempting to develop workable approaches and best-practice preservation strategies. Some internationally prominent example are Open Archival Information System (OAIS) by Consultative Committee for Space in United States, the PANDORA and PADI (Preserving Access to Digital Information) project by National Library of Australia (Webb, 2000)

The Arts & Humanities Data Service (AHDS) of the United Kingdom in its document "Digital   Information:a   guide   to   web   resources"   (viewed   at http://ahds.ac.uk/resource/preserve.htm#studies) brings out information about the large-scale preservation initiatives and projects and onto sites providing information on digital preservation initiatives, standards, technologies etc. A few major initiatives are as follows:

1. CEDARS (CURL Exemplars for Digital Archives) UK
    http://www.curl.ac.uk/projects/cedars.html
2. Commission on Preservation and Access (US)
     http://www.clir.org/cpa/)
3. Digital Library Federation (DLF, US)
    http://www.clir.org/diglib/preserve.htm)
4. International Federation of Library Associations (IFLA)
    http://www.ifla.org
5. International Standards Organisation (ISO), Open Archival Information Systems
    (OAIS Standard) http://ssdoo.gsfc.nasa.gov/nost/isoas/dads/)
6. Library of Congress (US) Digital Preservation Project
    http://lcweb.loc.gov/preserv/
7. Preserving Access to Digital Information (PADI, Australia)
    http://www.nla.gov.au/padi/
9. Research Libraries Group (RLG, US)
    http://www.rlg.org./
10. National Library of Canada Electronic Publications Pilot Project
     http://collection.nlc-bnc.ca/e-coll-e/ereport.htm

4.   Conclusion

Many libraries are still experimenting with the existing and various other options for access to material provided by electronic means. The future of research and scholarship depends on the ability to preserve digital resources into the future. The preservation and long term access to digital materials will be an era of concern for libraries and other organizations involved in the preservation of our scholarly and cultural heritage well into the new millennium. Although the technical challenges are great, there are lot of other non-technical issues that will need to be addressed like : building up a legal framework and bringing out consensus on widely accepted standards relating to digital preservation. Libraries should have strategic plans and policies towards building the technological infrastructure and procedures, establishing mechanisms of recording metadata, moving digital resources from less stable carriers to more stable ones and identifying formats for easier migration of digital information.

5.   References

1. Conway, Paul. Digitizing preservation. *Library Journal*, February; 1994; 42-45

2. Denise, Lievesley. Strategies for Managing Electronic Resources. In: Long Term Preservation of Electronic Materials. A JISC British Library Workshop. 27-28th November 1995 (Available at: http://www.ukoln.ac.uk/)

3. Russel, Kelly. Digital preservation: ensuring access to digital materials into the future. 1999 (Available at: www.leeds.ac.uk/cedars/chapter.htm)

4. Fresco M. Digital preservation: state of the art. Conference Proceedings of DG Information Society, Luxembourg, 17th-19th November 1999

5. Commission of Preservation and Access and Research Libraries Group, Inc. Preserving digital information: report of the Task Force on Archiving of Digital Information. 1999. (Available at: www.rlg.org./ArchTF/tfadi.index.htm)

6. Hendley, Tony. Comparison of methods and costs of digital preservation: British Library Research and Innovation report,106; 1998. (Available at http://ww.ukoln.ac.uk/services/elib/papers/tavistock/)

7. Hedstrom, Margaret. Digital preservation: matching problems, requirements and solutions (Available at: http://www.cess.missouri.edu/NSF Workshop/hedpp.html)

8. Webb, Colin. The role of preservation and the library of the future. 11th Congress of South East Asian Librarians Conference (CONSAL), Singapore, 26-28 April, 2000

9. Hedstrom, Margaret & Montgomery, 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)

10. Rothenberg, Jeff. Avoiding technological quickstand: finding a viable technical foundation for digital preservation; 1999 (Available at: http://www.clir.org/pubs/reports/rothenberg/contents.html)