Digital Preservation: Major Issues, Challenges and Possible Solutions in Digital Environment

B Ramesh Babu¹ and Ramesha²

¹Department of Information Science, University of Madras, Chennai
E-mail: beeraka_r@yahoo.co.uk
²Department of Library and Information Science, Bangalore University, Bangalore.
E-mail: bramesha@rediffmail.com

1. INTRODUCTION

Libraries and archives play a critical role in organizing, preserving, and providing access to the cultural and historical heritage of the society. In the relatively stable world of printing, hand-written and mechanically-reproduced information, libraries and archives managed to preserve this rich heritage for specialized scholars and for the general public. Libraries and archives are serving as the central institutional focii for preservation and both types of institutions include preservation as one of their core functions. In recent decades, many major libraries and archives have established formal preservation programmes for traditional as well as digital resources. The introduction of digital technologies into the process of production, distribution, and storage of information challenges the capacity of libraries and archives to carry out their responsibilities for preservation. Information management covers the whole spectrum of information handling activities, technology, covering methods of inputting, storing, retrieving, and distributing information. Its function incorporates a wide range of disparate activities including records management, library management, printing, reprography and micrography. The present day knowledge society has witnessed the widespread transition of knowledge from print format to electronic format as well as online, which has also given rise to the problem of its preservation in digital form. The problem of preservation is further complicated by the rapid obsolescence of the hardware and software. Ensuring continued access to digital information necessarily involves copying or transforming digital documents to run on current media, software, hardware and operating systems.

2. LIBRARIANSHIP

Four wheels of librarianship are: Collection of materials appropriate for libraries, organisation of the collection, preservation of materials, dissemination of information.

2.1 Preservation

As per the definition given by International Federation of Library Association (IFLA), Preservation includes all the managerial and financial considerations, including storage and accommodation provisions, stabbing levels, policies, techniques and methods involved in preserving library and archive materials and the information contained in them.

As per Conway¹ the term preservation is an umbrella under which most librarians
chivists cluster all of the policies and options for action, including conservation efforts. It has been the responsibility of librarians and archivists—and the clerks who went before them—to assemble and organize documentation of activity in places where it can be protected and used.

**Conservation**

According to IFLA Conservation Denotes those specific policies and practices devised in protecting library and archive materials from deterioration damage and including the methods and techniques devised by technical staff.

**Restoration**

As per the definition given by IFLA, restoration denotes those techniques and methods used by technical staff engaged in the making good of library and archive materials damaged by time / use and other factors.

**Digital Resources**

The creation of digital libraries presupposes the creation of electronic or digital resources and conversion or digitization of the existing resources into digital forms. In simplest connotation, digital resources refer to any source, which is in digitized form that can be read and scanned by means of electronic media. Unlike conventional resources, digital resources do not require separate space in a library, as these can be in a computer locally or remotely. Digital resources include a wide range of types such as:

- Collections in which complete contents of documents are created or converted to machine readable form for online access;
- Canned images, images of photographic or printed texts etc;
- Online databases and CD-ROM information products;
- Computer storage devices such as optical discs, juke boxes, CD ROM/DVD-ROM;
- Databases accessible through Internet and other networks;
- Meta documents;
- Digital audio, video clips or full-length movies.

**Loss of Digital Information**

As we move into the electronic era of digital objects, it is important to know that what we know today, much of what is coded and written electronically, will forever. Many digital resources frequently change their content, presentation; mechanisms need to be developed for digital libraries to keep track changes and of different versions of the same resources.

Some of the reasons for the loss of digital information are:

- Changes in an organization
- Content re-organization
- Dissolution of sponsorship
- Technology obsolescence
- Content format obsolescence
- Hacking and sabotage
- Disasters (natural or man-made)
5. **DIGITAL PRESERVATION**

Digital preservation means taking steps to ensure the longevity of the electronic document in terms of the following:

- Data (text, image, video or audio stored in a variety of formats and standards)
- Index to the data
- Link to other data
- Metadata
- Software (relies upon hardware and operating system) and
- Storage media, etc.

Digital preservation concerns two types of documents: namely ‘born-digital documents’ and ‘digitally created’ documents.

**Born digital documents**: These refer to those materials that were initially created using some form of digital technology. They are often termed as ‘electronic records’.

**Digitally created documents**: These refer to those materials, which have been transformed from analog to digital form through some reproductive means such as rekeying the information or scanning the document or object etc.

The preservation of digital objects involves a variety of challenges, including policy questions, institutional roles and relationships, legal issues, intellectual property rights, and metadata. In addition to these issues, there are following substantial challenges at the empirical level.

- What does it mean to preserve digital objects?
- What purposes are served by preserving them?
- What are the real possibilities for successful preservation?
- What are the problems encountered in trying to exploit these possibilities?
- Can we articulate a framework or an overall architecture for digital preservation that allows us to discriminate and select possibilities?

Ensuring the long-term preservation of information in digital form will be one of the greatest challenges for the information professionals in the 21st century. While there has been an awareness of digital preservation problems for some time, their importance have recently been magnified because of the increasing dependence of the world on computers and networks.

### 5.1 Digital Preservation Tasks

According to Hendley, at a basic level, all digital preservation strategies involve the following tasks:

- **Preservation of digital medium** that holds digital information by storing it in the correct environment and following agreed storage and handling procedures.
- **Preservation of the technology** by conserving the hardware and software environment used to create and use the resource.
- **Preservation of intellectual content** by preserving the integrity of the digital information during the copying process.

### 5.2 Physical Storage Conditions Required for Digital Objects

The digital objects should be stored in a functional room, which provide thermal comfort so that constant temperature and humidity can be maintained. The Temperature required is as low as 5°C and Relative Humidity as low as 20 per cent, Temperature variations not to exceed 4°C and humidity variations not to exceed 10 per cent RH. Ideally, light would be artificial and if there are windows
storage room, then the north-facing position is best to avoid direct sunlight. Similarly, the storeroom should have diffuse lighting. The digital objects should not be stored from water, heat, and strong magnetic fields. The digital objects should be stored in a vertical position in appropriate containers and extra care should be taken for effective handling of tapes, cassettes, or disks, etc. The environment should be kept clean from dust and smoke (pollution-free). To reduce wear and tear, an access copy to be made for frequently used materials and to the playback equipments clean and in good running condition. Whenever problems arise while handling the digital objects, professionally trained personnel should be asked to tackle the problem. This would minimize problem in handling information resources.

CONSIDERATIONS IN DIGITAL PRESERVATION

In the task of digital preservation, some of the aspects are to be considered are the object, physical preservation, content preservation, presentation format, reliability, authenticity, location and references, and provenance.

In preserving a digital object, one should aim at:

- **retain the object as a discrete whole:** A book is a book, or it exists in a set, but what are the boundaries of a document in a hypertext environment? The boundaries of digital objects are less clear, especially if they are compound objects created by assembling different media or by linking to resources from around a network.

- **preserve the physical presence:** The physical presence in this case refers to the computer file, the series of “1”s and “0”s that are the basis of a digital object. As we shall see, preserving the physical file does not mean that the object will remain accessible.

- **retain content:** This aspect refers to maintaining the ability to access the content at its lowest level, such as ASCII text, without the embellishments of font variations and layout features.

- **preserve the presentation:** Content is typically rendered in some presentation, format or layout that includes different font faces and sizes, the use of white space, columns, marginalia, headers, footers, pagination, and so on. In many types of digital documents (e.g., SGML, XML, and some forms of PDF), the layout specifications are separate from the content. To retain the original look of a document, these layout specifications must also be preserved, especially when they contribute significantly to the understanding and interpretation of the content.

- **preserve functionality:** Digital objects have a functionality that goes far beyond additional paper documents. They can contain multimedia components (i.e., text, graphics, audio, and video), exist in hypertext format, contain dynamic content generated automatically from data stores, or have navigation functions, such as toolbars, keyword search, or interactive tables of contents. Special efforts must be made to preserve the functionality.

- **preserve authenticity:** An individual accessing the object must be able to verify that it is the one who wants, and that the transformations to keep it accessible have preserved its original form. Thus, activities to guard authenticity include securing digital objects against unauthorized changes and monitoring digital objects through multiple “copying” cycles to ensure that each copy is an accurate representation of the original.
acceptable rendition of the original. Establishing authoritative depositories would also help achieve this goal. Ensuring the authenticity of digital information sources is a complex and resource-intensive process. Rothenberg proposed a number of strategies that may be considered to ensure the authenticity of digital information resources. These are:

- **Originality strategy**: to assess whether the information has altered from its original state.
- **Intrinsic properties strategies**: to identify certain properties of an informational entity that define authenticity, regardless of whether they imply the originality of the entity.
- **Suitability strategy**: to define authenticity in terms of whether an informational entity is suitable for some purpose.
- **Locate and refer to the digital object over time**: Digital objects can be readily altered, copied or moved. An individual must be able to match a citation to a digital object, and to distinguish it from other versions or editions.
- **Preserve provenance**: Provenance is an archival concept that asserts the origin and chain of custody of an object and contributes to defining it as a whole. Imprint statements and bookplates, for example, partly fill this role for formally published items. Establishing an object's origin and history help confirm that the work is authentic and its content are intact.
- **Preserve context**: Digital objects are partly defined by their hardware and software dependencies, their mode of distribution and linkages to other digital objects. Preserving digital objects may mean weaning them from some technical dependencies, changing the mode of distribution, and deactivating links. In these circumstances, preserving context is a particular challenge.

7. **ISSUES RELATED TO DIGITAL PRESERVATION**

Preservation is a response to the threat of destruction. Information professionals all over the world are trying to rejuvenate their professional practices to ensure long-term preservation of digital information. Day identified the following issues related to digital preservation, which are worthy to consider:

- **Preservation strategies**: Strategies concerned with digital preservation need to be involved at the beginning of the digital life cycle.
- **Data creators and publishers**: It is vital that there is a sound and effective dialogue between creators and publishers of digital information and preservation professionals.
- **Intellectual Property Rights**: Preservation rights need to be part of license negotiations between publishers and library consortia.
- **Collection management**: There shall be Collection Development Policy (CDP) for digital information to maintain the possibility of future serendipity.
- **Metadata**: Need to develop appropriate standards on Metadata for digital preservation.
- **Staff expertise**: The education and training programmes in LIS shall focus on the staff skills and expertise in digital preservation.

There is no shortage of issues related to digital preservation and as Hunter remarks that information professionals owe it to their institutions to confront the digital preservation questions head-on. They may prove to be some of the most difficult professional challenges over the next few decades.
DIGITAL PRESERVATION: IMPACT ON LIS TRAINING AND EDUCATION

Recent developments in the field of digitization and digital preservation have created challenges in the LIS education. To cope with the changes in this field, there is an urgent need to train the professionals as well as students to enable them to face the challenges in the light of digitalization activities in the libraries. To this end, it is necessary to redesign the course contents and add a new course to the existing curriculum that can be considered for the education and training of prospective information professionals in the context of digital preservation. The introduction of Preservation of Information Materials in the Digital Era is a course that could be considered for education and training purposes. The contents of this course would include the following modules:

- Introduction to Preservation and Conservation and related concepts
- Evolution of Information materials from clay tablets to digital forms
- Nature and types of Information materials – Detailed study
- Hazards to Information materials covering environmental; physical and chemical, biological, and disasters, etc.
- Physical and chemical methods of treatment involving fumigation, lamination, encapsulation, binding and mending, restoration, leaf casting, etc.
- Preservation and conservation of print, digital and non-book materials and building and equipments.
- Role of micrography and reprography in preservation.
- Digitization covering means and methods; practical and strategic issues in the digitization of information materials; an overview of major digitization projects.
- Digital preservation issues covering trends, challenges and strategies, national and international initiatives in digital preservation.

The above-stated course modules can be considered as a core or an elective in the LIS School as the case may be. In addition to the formal training on object, the LIS schools may also consider organizing Continuing Education programs or In-service training programs such as short-term courses, refresher sessions and workshops in this area. This would benefit the in-service personnel, teachers and practicing LIS professionals.

DIGITAL PRESERVATION: SOME SUGGESTIONS

Keeping in view of the above factors, some of the suggestions are:

- Draft a set of guidelines, standards and practices for digital preservation.
- Develop means to coordinate digital preservation activities among institutions.
- Develop institutional policies for acquisition, conversion, storage and maintenance of digital materials.
- Redesign Library and Information Science curriculum with reference to digital preservation and digitization.
- Plan and programme for education and training of digital preservationists.
- Prepare a model for infrastructure required for digital preservation.
- Sharing of technical know-how of digital preservation among the fellow professionals.

CONCLUSIONS

Preservation of digital information is complex because of the dependency digital nation has on its technical environment. Digital preservation encompasses a range of activities designed to extend the usable life of machine-readable...
computer files and protect these from media failure, physical loss, and obsolescence. Digital preservation will add little values to the research process if it serves only as an alternative form of storage. Preserving digital materials in formats that are reliable and usable will require long-term maintenance of structural characteristics, descriptive metadata, display and computational and analytical capabilities which demand mass storage and software for retrieval and interpretation. The digital preservation is a process that requires the use of the best available technology, careful thought, administrative policy and procedures.

Digital preservation policies and practices are not currently well developed in Indian libraries, and only a few library and information professionals have assumed responsibility for preserving materials in digital form. Preservation in the area of digital technology is a shared responsibility. The wide community of information professionals in libraries, archives, and museums has collaborated in developing strategies to build technically sustainable solutions. There is an ongoing need to work as a community to drive the development of needed products, to grow a demand for commercial imaging services, and support for international standards to meet digital preservation needs. Growing from the initiative of a national cooperative digital imaging project, a further need for coordination among the various parties involved in digital preservation to take concrete measures to build capacity and increase staff expertise with issues of digital technology. There is an inherent paradox in digital preservation. On the one hand, it aims to deliver the past to the future in an unaltered, authentic state. On the other hand, doing so inevitably requires some alteration.

The Archives, Libraries, Museums and other Repositories, have been instrumental in preserving and providing access to scholarly communications, documentary heritage, and other cultural resources in traditional formats. It is clear that many of these institutions are beginning to adapt digital preservation to their array of preservation responsibilities. There are no digital preservation policies or established methods to preserve digital information. Digital preservation is a new challenge and they are just beginning to confront the policy, technological and human resource implications. In digital preservation, all that looks new is old again. Keeping in view the recent developments in the information handling and management, the benefits of digital preservation and digitization could not be overlooked and augmented effectively to preserve and disseminate for the future generations to come. Therefore, the information professionals are to be trained in the area of digital preservation and digitization techniques.

The emerging research on preserving both born-digital and born-again-digital documents is vibrant and working toward the development of strategies, technologies, and methods that will enable digital preservation to become an established reality. To conclude "the field of preservation is moving forwards rapidly in a great many areas as the profession seeks simultaneously to raise the consciousness of librarians and the public, to provide better education and training, to stimulate the development of programmes at all levels and to define this still embryonic field" as stated by Lisa Fox.

REFERENCES