Digital Preservation for Libraries: Planning / Process and Approaches

V.R. Rajan*, S. Barathi**, D. Sindhujha**

Abstract

This paper discusses about the digital preservation – planning, process and approaches. It also deals with the major part process, issues and technologies involved in preservation of digital materials, explains the overview of technological approaches and strategies to digital preservation and challenges,focuses on operational digital preservation systems specifically in information resource centre (IRC).It considers the wide range of digital objects of interest to IRC, including e-journals, technical reports, e-records, project documents, scientific data, etc. The paper also discusses archiving based on format types – text, data, audio, video, etc. Design, methodology, approach. And also find out that the Digital preservation of documents restores it from loss, theft and decay.

Keywords: Digital Preservation; Preservation Planning; Digital Resources; Preservation Initiatives; Preservation Strategies.

Introduction

"The Goal of Any Preservation Program is to ensure long – term, ready access to the Information Resources of an Institution."

Abby Smith "Preservation in the Future Tense"

Digital preservation can be defined as the combination of policies, strategies and actions to ensure access to and accurate rendering of authenticated reformatted and born digital content over time, regardless of the challenges of media failure and technological change. Digital preservation is a general term for a set of services with one goal: keeping your data safe so that it can be accessed in the future. Digital preservation is the management of digital information over time. It takes the form of processes and activities that ensure continued access

to information and all kinds of records, both scientific and cultural heritage, that exists in digital form.It is also about the preservation of digital assets and their associated metadata.As Marcum noted in 1996, "Preservation is a fundamental responsibility of libraries and archives of record", this paper discusses Digital preservation aims to ensure that a digital collection remains usable, regardless of the inevitable changes in technology the future will bring. Without the appropriate preservation methods in place a digital collection can easily become inaccessible and so useless in just a few years. 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

Purpose of Preservation

To preserve a complete record of all important information, so that data has meaning in the future, ensuring that it meets the requirements of funder mandates, audits, restoration of missing content and re-use of data. Preservation systems retain data for many years.

Digital Preservaton (Digital Documents Preservation)

- The aim of digital preservation is long-term, error-free storage of digital information, with the means

Author's Affiliation: *Deputy Librarian (Retd.), Pondicherry University, Kalapet, Puducherry, 605014, India. ** Ph.D Research Scholar (LIS), Amet University, Chennai, Tamil Nadu 603112, India.

Reprint's Request: V.R. Rajan, Deputy Librarian (Retd.), Pondicherry University, Kalapet, Puducherry, 605014, India. E-mail: rajan_arthi@yahoo.com

of retrieval and interpretation, for the period of time that information is required

- Digital preservation refers to the series of managed activities necessary to ensure continued access to digital materials for as long as necessary

- *Long-Term Preservation*

Continued access to digital materials, or at least to the information contained in them, indefinitely

- *Medium-Term Preservation*

Continued access to digital materials beyond changes in technology for a defined period of time but not indefinitely

- *Short-Term Preservation*

Access to digital materials either for a defined period of time while use is predicted but which does not extend beyond the foreseeable future and/or until it becomes inaccessible because of changes in technology

Digital Preservation – Needs

The digital world is a place of rapid technological and organizational changes, which impacts on the continuing use of digital resources. All types of digital resources need preservation. Some of the digital resources listed below:
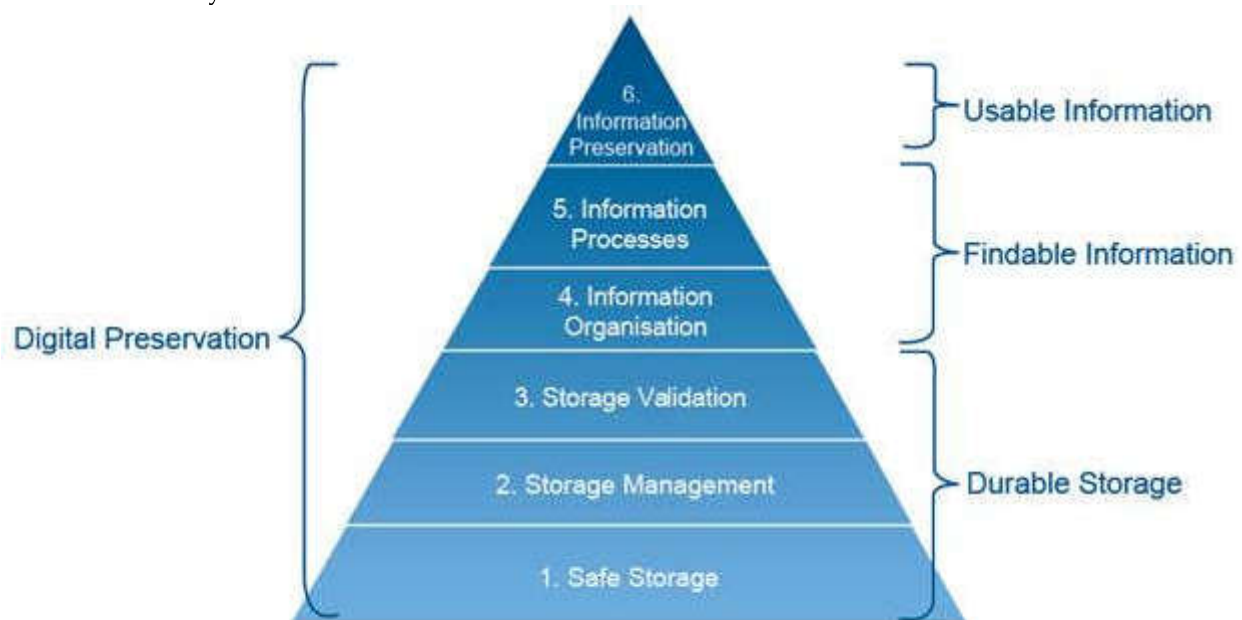
Digitally Reformatted

Digitized versions or surrogates of physical items

Born Digital

Digital resources that have no analogue counterpart

Individual resources

Texts, still and moving images, sound recordings, etc.

Collective resources

Websites, e-journals, wikis, catalogues, etc

Data Sets

Scientific and cultural data comprising multiple individual pieces of data

Communication records

Email, instant messages, etc

Digital Preservation

Contains a wide range of information, with links to tutorials, videos, blogs, to help provide some basic context. It also includes an overall definition of digital preservation: the series of managed activities necessary to ensure meaningful continued access, for as long as it is required, to digital objects and materials.

Preservation by Format

Includes links to suggested approaches for preserving photographs, audio, video, email, documents, and websites. Much of this is focused on smaller, personal collections.

Digital Storage, Cloud computing and Personal Backup

Includes links to basic information on cloud storage and other backup options, as well as a timeline history of digital storage

Process of Digital Preservation

Digital Assets

• Creating & shielding Assets

• Accessing & Discovering Assets

• Managing & Preserving Assets

• Understanding & Awareness of Assets

Digital Preservation: A Four Step Process

To preserve digital records, the National Archives uses its digital preservation software to convert proprietary file formats to openly-specified preservation file formats.

There Are Four Steps in this Process

1. Manifest – A list is created of all digital records to be preserved.

2. Quarantine – Records are checked for viruses and integrity.

3. Preservation – Records are converted to preservation file formats.

4. Storage – Records are stored in the digital archive.

The Archives Digital Preservation Software

• Captures the essential elements of digital records

• Allows digital records to be retrieved from the digital archive at any time

• Continually checks the integrity of records in the digital archive.

Approaches to Preservation

The protection and long-term preservation of a project's digital output and metadata needs to be considered even before digitization begins. Decisions made at the earliest stages of a project can and will have an impact on the effectiveness of the whole digital preservation strategy. It is particularly important that digitization projects are fully documented as they progress. Full documentation of technical solutions and project delivery will give those undertaking the preservation strategy an understanding of how the project was conceived, developed and produced.

Strategies have to be put in place to guarantee that the collection survives through technological changes, ensuring its continued accessibility and usability. There are three common approaches to digital preservation:

• *Migration:* Itdescribes the process of copying content from one format (such as a CD-ROM) onto a newer format (such as a solid state flash drive).

• *Refreshment:* A related process is refreshment. Refreshment involves copying data onto a newer example of the same format (such as from an old CD-ROM to a new CD-ROM).

• *Emulation:* Emulation is a more involved process of accessing data on a system other than the one it was made for. Commonly, this will be because an original system is no longer available. Playing vintage computer games on a contemporary games emulator is a good example.

Preservation in Practice

Whichever approach or combination of approaches is chosen, it is often helpful to make a distinction between a 'master generation' of digital data and at least one surrogate 'delivery generation'. The master generation should contain as much intellectual, visual or audio content as possible and must be saved in a standard (non-proprietary) file format and it should preferably be duplicated across multiple locations. Delivery generations of data, however, may be re-sized, compressed, and saved in whichever format is suitable for delivery to the user. Delivery versions are typically of lower quality (more compressed) than their original master files. Defining the status and thereby the relative importance of a file helps immensely in the task of preservation

Digital Preservation Strargies

Issues

Digital preservation encompasses a range of strategies, processes and activities, with a variety of associated issues to be considered. Examples are:

Long-term: May extend indefinitely and depends on the need for continuing access to a resource in one or more specific formats. The lifetime of a specific resource is determined by the degradation and/or format accessibility of that resource.

Retrieval: Obtaining digital files from storage without corrupting the stored files.

Interpretation: The digital files must be decoded and

transformed into usable representations, for machine processing and/or human access.

Rendering: Making a digital file available for a human to access.

Re-digitizing: Some early digitized resources are in formats that are, or are rapidly becoming, obsolete. Since it can be the case that poor results are obtained by migrating from the obsolete format to a newer format, it may sometimes be better to re-digitize from the original.

Emulation: Where specific playback equipment is no longer available, emulation software may need to be written in order to access the informational content using a different device.

Degradation: The process by which parts of a resource are lost over time. This may occur as a characteristic of a format (it becomes a less accurate representation over time) or a consequence of copying from another file or migrating from one format to another.

Effort: It appears that digital preservation requires more frequent and ongoing action than other types of media. The consequent requirement in terms of effort, time and money

Technical Strategies

Digital preservation to date has relied on two main technical strategies: standards and migration. Technical standards form a foundation for much of what makes digital libraries possible. Standards and protocols for storage, data formats, bibliographic control, display, retrieval, transport, and distribution are imbedded in the infrastructure that make digital libraries accessible, manageable, and useable. In the area of digital preservation, standards issues primarily concern encoding, data formats, and representation schemes. Archivists and librarians tend to favor open standards over *de facto* or proprietary standards for two reasons. First, open standards are published and readily available whereas *de facto* standards

Operational Models and Programs

Digital preservation is difficult to untangle from many other digital library functions. Preservation concerns may affect collection development if a digital library limits the materials it will acquire to those that conform to designated standards. Likewise, preserving digital materials would be pointless if a digital library could not provide a means for accessing the materials. The inter-relationships between

collection development, preservation, and access, however, add levels of complexity to digital library development that are difficult to address without some method of breaking the inter-dependencies down into reasonable and solvable problems. These closely related issues also influence how digital libraries are designed and administered. Recent conceptual work on archival functions and efforts to develop operational digital repositories provide some useful frames of reference for integrating digital preservation into digital library design and operations.

Current Research and Ongoing Issues

The increasing amount of research underway that directly or indirectly addresses concerns about the longevity of information is another encouraging sign that digital archiving has become an important issue for digital libraries.

- A number of open source products have been developed to assist with digital preservation, including Archivematica, DSpace, Fedora Commons, OPUS, SobekCM and EPrints. The commercial sector also offers digital preservation software tools, such as Ex Libris Ltd.'s *Rosetta*, Preservica's Cloud, Standard and Enterprise Editions, CONTENTdm, Digital Commons, Equella, intraLibrary, Open Repository and Vita.

Examples of Digital Preservation Initiatives

- The Library of Congress founded the National Digital Stewardship Alliance which is now hosted by the Digital Library Federation.

- The British Library is responsible for several programmes in the area of digital preservation and is a founding member of the Digital Preservation Coalition and Open Preservation Foundation. Their digital preservation strategy is publicly available. The National Archives of the United Kingdom have also pioneered various initiatives in the field of digital preservation.

Conclusion

Recent progress in digital preservation is a consequence of the growing awareness of longevity as a critical issue for sustainable and useable digital libraries, increased investments in research and development, and efforts to focus on discrete and potentially solvable aspects of the problem. These

issues will provide ample opportunities for research, experimentation, and development for years to come.This paper has described a methodology for expanding existing offerings and building new ones. While these offerings will undergo transformation, we are building them with the certainty that users in centuries to come will find our early collaborative efforts in digital preservation to have been valuable.

References

1. Digital preservation Coalition
2. The UC SAN DIEGO
3. Digital preservation Management
4. *DESIDOC Journal of Library & Information Technology*
5. *e-Library Science Research Journal*
6. Digital Preservation-Wickepedia