The Preservation of Digital Information: Issues and Observations

Sheela V.

ABSTRACT

Preservation of digital information is widely considered to require more constant and ongoing attention than preservation of other media. A library is an organization. The organization of libraries is already changing as electronic information increasingly becomes part of their charge. Digital imaging technology offers distinctive advantages to institutions with impressive collections of scholarly resources. Information content can be delivered directly to the reader without human intervention by readers remotely, although such delivery may tax the capabilities of even the most sophisticated projection equipment and networks. This article explains the different types of digital resources in general and digital preservation in particular based on the user's benefits.

Keywords: Information Dissemination, Digital Resources, Digital Preservation

INTRODUCTION

Digital Preservation is the management of digital information over time. The constant input of effort, time and money to handle rapid technological and organizational advance is considered the main stumbling block for preserving digital information. Indeed, while we are still able to read our written heritage from several thousands years ago, now the digital information created more demands. As libraries move more into the electronic environment. In addition, the need for consortial activity has become evident both for provision and preservation of digital information. Digital collections, however, as a proposition of the libraries supply of information, will grow for the foreseeable future and the quantity of information requiring are will become considerable. The advent of electronic information as libraries takes the responsibility for the preservation of information in non-artifactual forms. Electronic information will have to make a conscious commitment to providing resources. Where the role of a library is to facilitate access to information ''Digital Preservation refers to the various methods of keeping digital materials alive into the future'' according to recent statement from the council on library and information resources. Digital preservation typically centers on the choice of interim storage media. The term digital preservation means the planning resource allocation and application of preservation methods and technologies necessary to ensure that digital information of contemning value remains accessible and usable. The concept of digital preservation encompasses material that begins its life in digital form as well as materials that are converted from non-digital to digital formats. Digital preservation can therefore be seen as the set of processes and activities that ensure continued access to information and all kinds of records, scientific and cultural heritage existing in digital formats. In the language of digital imaging and electronic resources, preservation is no longer just the product of a program but an ongoing process. In this regard the way digital information is stored is

Aythor's Affiliations: Assistant Librarian, Karnataka State Open University, Manasagangotri, Mysore–06, Karnataka

Reprints request: Sheela.V, Assistant Librarian, Karnataka State Open University, Manasagangotri, Mysore– 06, Karnataka

E-mail: she579@gmail.com, Phone No: (M): 9341024224, (R): 0821-2344615.

important in ensuring their longevity. The long –term, error–free storage of digital information, with means for retrieval and interpretation, for the entire time span the information is required for. In this ''long-term'' is defined as'' long enough to be concerned with the impacts of changing technologies, including support for new media and data formats or with a changing user community. ''retrieval'' means obtaining needed digital files from the long term, error – free digital storage, without files from the long term, error –free digital storage, without possibility of corrupting the continued error – free storage of digestion files. ''Interpretation means that the retrieved digital files are decoded and transformed into usable representations.

OBJECTIVE OF THE ACTIVE PRESERVATION

Now a large quantity of information exists in digital forms, including e mails, blogs, social networking website, national electronics website, web photo albums and sites which change their content over time. The unique characteristic of digital forms makes it easy to create content and keep it up –to- date, but at the same time brings many difficulties in the preservation of this content. Margaret Hedstrom points out that'' digital preservation raises challenges of a fundamentally different nature which are added to the problems of preserving traditional format materials. In order to understand the tasks of digital preservation, we need to device some working definitions for the concepts of ''document'' and ''digital objective'' (Levy must be in bibliographic reference) David Levy has offered some useful intuitive definitions ''Documents are talking things, they are bits of the material world'' as paper and printing technologies have natured, people have grown to expect a document to hold human verbal communication fixed so that it can be repeated. The great advantage of digital media, the case of copying and modification, also becomes a major liability.

The RLG report state that 'A reliable digital repository is one whose mission is to provide long–term access to managed digital resources, that accepts responsibility for the long term maintenance of digital resources on behalf of its depositors and for the benefit of current and future users. Then the finally digital preservation refers to the various methods of keeping digital materials alive into the future. The most common application of digital technologies is an library is digital copies that can be used for ready reference in live of casual browsing through the original sources. Preservation goals are met because physical access to the original document is limited. Additionally, digital products that draw together, organize and enhance access to widely dispersed research materials may have transcendental impact on the people who use them. Each of these preservation applications places separate but increasingly rigorous demands on digital technologies.

CONVERSION OF INFORMATION FROM ANALOG TO DIGITAL

A major challenge in choosing paths from analog to digital is obtaining an in-depth understanding of the particular characteristics of the collections or the individual items being converted. The most important characteristics are:

- Format of the source including size, its structure and its physical condition.

- Physical condition and its impact on the ability of the item to be handled during the conversion process.

- Visual characteristics

- Colour as an essential carrier of information content

- Level of detail

Beyond these specific characteristics there is a significant impact on the cost, quality of the conversion project.

CHALLENGES OF DIGITAL PRESERVATION

The first challenge of digital preservation faces is that the media on which digital contents stand

are more vulnerable to deterioration and catastrophic loss. While acid paper is prone to deterioration in terms of brittleness and yellowness, the deterioration really happens it happens over decades too. It is also highly possible to retrieve all information without loss after deterioration is spotted. Once the deterioration starts is most cases there is already data loss. This characteristic of digital forms leaves a very short time frame for preservation decisions and actions. Another challenge is the absence of established standards, protocols and proven methods for preserving digital information. Another challenge, perhaps a more serious and important one is the problem of long-term access. Digital technology is developing extremely fast and one retrieval and playback technology can become absolute in a matter in a year. When faster, more capable and cheaper storage and processing devices are developed the older version gets replaced almost immediately. This process is known as digital obsolescence.

This challenge is exacerbated by the lack of established standards, protocols and proven methods for preserving digital information. We used to save copies of data on tapes, but media standards for tapes have changed considerably over the last five to ten years and there is no guarantee that tapes will be readable in the future.

Hedstrom further explained almost all digital library researches have been focused on 'architectures and systems for information organization and retrieval, presentation and visualization and administration of intellectual property rights' and that 'digital preservation remains largely experimental and replete with the risks associated with untested methods'. While the rapid advance of technology threats access of digital contents in length.

STRATEGIES OF DIGITAL PRESERVATION

In 2006, the online computer library center (OCLC) developed a four point strategy for the long-term preservation of digital objects that consisted of:

1. Assessing the risks for loss of content posed by technology variable such as commonly used proprietary file formats and software applications.

2. Evaluating the digital content objects to determine what type and degree of format conversion or other preservation actions should be applied.

3. Determining the appropriate metadata needed for each object type and show it is associated with the objects.

4. Providing access to the content.

There are several additional strategies that individuals and organizations may use to actively combat the loss of digital information.

Refreshing is the transfer of data between two types of the same storage medium so there are no bitrate changes or atteration of data. Refreshing will likely always be necessary due to the deterioration of physical media.

Migration is the transferring of data to newer system environments. This may include conversion of resources from one format to another. Resources that are migrated run the risk of losing some type of functionality since newer formats may be incapable of capturing all the functionality of the original format, or the converter itself may be unable to interpret all the nuances of the original format.

The National Archives electronic records archives and Lockheed Martin are jointly developing a migration system that will preserve any type of document, created on any application or platform and delivered to the archives on any type of digital media.

Creating duplicate copies of data on one or more systems is called replication. Data that exists as a single copy in only one location is highly vulnerable to software or hardware failure, intentional or accidental atteration and environmental catastrophes like fire, flooding etc. digital data is more likely to survive if it is replicated in several locations.

Emulation is the replicating of functionality of an absolute system. Emulation has been a popular strategy for retaining the functionality of old video game systems. The feasibility of

emulsion as a catch –all solution has been debated in the academic community.

Metadata attachment is data on a digital file that includes information on creation, access rights, restrictions, preservation history and rights management, metadata attached to digital files may be affected by file format obsolescence. It retains information, but not the structure information is presented in.

Digital objects that can speak to their own authenticity are called trustworthy digital objects it enable digital objects to maintain a record of their change history so future users can know with certainty that the contents of the object are authentic.

Digital sustainability encompasses a range of issues and concerns that contribute to the longevity if digital information. Unlike traditional, temporary strategies and more permanent solutions, digital sustainability implies a more active and continuous process, digital sustainability incorporates activities in the present that will facilitate access and availability in the future.

STANDARDIZATION OF DIGITAL PRESERVATION

To standardize digital preservation practice and provide a set of recommendations for preservation program implementation, the Reference Model for an Open Archival Information System (OAIS) was developed.

An OAIS is an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a designated community. The information being maintained has been deemed to need Long Term preservation. Long Term may extend indefinitely. In this reference model there is a particular focus on digital information both as the primary forms of information held and as supporting information for both digitally and physically archived materials. This reference model:

- Provides a framework for the understanding and increased awareness of archival concepts needed for Long Term digital information preservation and access.

- Provides the concepts needed by non – archival organization to be effective participation in the preservation process.

- Provides a framework, including terminology & concepts, for describing and comparing architectures and operations of existing and future archives

- Provide a framework for describing and comparing different Long Term preservation strategies and techniques.

- Provides a basis for comparing the data models of digital information preserved by archives and for discussing how data models and the underlying information may change over time.

- Provides a foundation that may be expanded by other effort to cover Long Term preservation of information that is NOT in digital form.

- Expands consensus on the elements and process for Long Term digital information preservation and access and promotes a larger market which vendors can support.
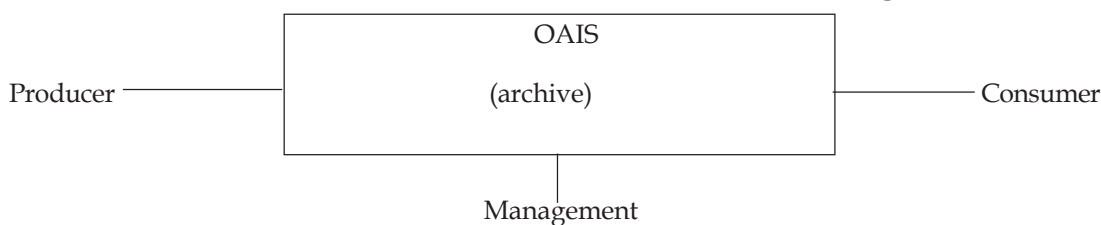
- Guides the identification and production of OAIS related standards.

The reference model addresses a full range of archival information preservation functions including ingest archival storage data management, access and dissemination. It also addresses the migration of digital information to new media and forms, the data models used to represent the information the role of software in information preservation and the exchange of digital information among achieves.

CONCLUSION

Digital information sources are often more expensive than print. The libraries responsibility to preserve electronic information is equal to its responsibilities for collections of printed materials and other formats. Attention should be given to electronic information in the development of a preservation plan for the library, the term digital resources is defined as any work encoded and made available for access through the use of a computer. It includes both online data and electronic data is physical

The simple model shown the environment surrounding an OAIS is

```
                    ┌─────────────────────────┐
                    │         OAIS            │
Producer ───────────┤                         ├──────────── Consumer
                    │       (archive)         │
                    └────────────┬────────────┘
                                 │
                           Management
```

formats. The criteria used to evaluate electronic resources do not greatly differ from those used for books or materials in other formats. As with the more traditional formats the cost of the work and the requirements of serving, cataloguing, storing and preserving must be considered in the decision. The library selects electronic works for its permanent collections which rank high on the following list of criteria; usefulness in serving the current or future informational needs of researches, reputation of the information provides, amount of unique information provided, scholarly content, currency of the information, frequency of updating and easy of access.

Academic libraries have a unique opportunity in the area of digital preservation. As libraries and other institutions embark on the digital preservation process, judgement must be used to balance risk against the maturity of the process. Documents that are extremely rare, we would like to say that we will preserve our cultural heritage materials in perpetuity. Indeed there are many challenges for doing digital preservation. There is much research yet to be done.

Finally, it must be said that there will always remain the element of trust in the organization that takes on the role of 'trusted digital resources'. We rarely doubt the accuracy and validity of our statement –although we want to have at least this same level of trust in digital resources. Ultimately, users will need to be able to trust the people and organizations who have taken on the responsibility for managing the processes and technology of digital preservation.

REFERENCE

1. Kusum Verma (ed). The electronic Library. Akansha publishing house. 2004.

2. Kaul. A.K. E-Libraries in 21st century. ABD publishers. 2008.

3. Garrett.J.D. et.al. Preserving digital information: Report of the task force on archiving of digital information. 1966.

4. Granger, Stewart. ''Emulation as a Digital Preservation Strategy'' D-Library Magazine. 2000; 6(10).

5. Edwards, Eli. Ephemeral to enduring : The Internet Archive and its Role in preserving Digital Media''. Information Technology & Libraries. 2004; 23(1).

6. Hoeven.J. ''Diosure: emulator for digital preservation: D-Library Magazine. 2007; 13(11/12).

7. Jantz. R & Giarlo. M.J. Digital preservation: Architecture and technology for trusted digital repositories''. D-Library Magzine. 2005; 11(6).

8. Ross.S. Changing Trains at Wigan: Digital preservation and the Future and Scholarship. London, National preservation office. 2005.

9. Suleman.H and Fox.E. A framework for building open digital libraries. D-Library Magazine, 2001; 7(12).

10. Jantz.R. Public opinion polls and digital preservation: An application of the Fedora Digital object repository system. D-Library Magazine. 2003; 9(11).

11. http://en.wikipedia.org/wiki/Digital_preservation http://public.ccds.org/publications/archive/650x0b1.pdf.