ON CHALLENGES OF DIGITAL PRESERVATION

Dr. Manish Kumar Singh,

Information Scientist,
Central Library,
B.H.U., Varanasi.

ABSTRACT

Need of Digital Preservation is identified by sectors of large libraries and archives, giving major impetus for digital preservation studies. Both, readers and publishers are moving from print-based journals to digital publication. Moreover, academic communications are increasingly being published in digital formats, including online education programs, discussion groups, bulletin boards, etc. There is also an increasing expenditure by libraries, both academic and public nature, in acquiring digital resources and converting physical resources to digital to improve access and limit contact with physical objects [1]. Apart from this, creation of digital records that needs to be archived for unforeseen requirements also leads to the need of preservation studies. All these preservation needs demand a technical, organizational, legal and economic approach to developing an infrastructure comprehensive enough to ensure preservation of digital assets in all relevant sectors [2]. The purpose of this work is to perform a critical observation of the challenges of digital preservation faced by the libraries and archives.

Keywords: *Digital preservation, Digital archival.*

INTRODUCTION

Preservation is done to ensure protection of information of enduring value for access by present and future generations. Large libraries and archives have provided the much needed institutional focus 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 programs for traditional materials which include regular allocation of resources for preservation, preventive measures to arrest deterioration of materials, remedial measures to restore the usability of selected materials, and the incorporation of preservation needs and requirements into overall program planning [3].

The need for preserving digital assets was always felt and becomes more pressing and keeps growing by the day by expansion of reositories. According to the International Data Corporation (IDC), by 2011 information that is created, captured, or replicated in digital form was expected to increase 10 fold over that produced in 2006. The compound annual growth rate between 2006 and 2011 was expected to be almost 60% [4].

DIGITAL PRESERVATION AND DIGITAL MATERIALS

The word preservation is derived from the Latin verb "praeservare" a composite term of two words: "prae", means before; and "servare", to keep safe. Thus, "praeservare" means "to keep the past things safe" [5]. However, digital preservation is broadly defined as "a series of actions and interventions required to ensure continued and reliable access to authentic digital objects for as long as they are deemed to be of value" [6]. On the other hand, it can also be defined as "the actions needed to assure enduring access to the full content of digital resources over time" [7].

The activity of preservation is a response to the threat of destruction and loss. Recognition that there is a threat, elicits a response, the scale of which is usually in proportion to the value that is placed on the object under threat. The phrase "digital preservation" refers or embraces two major categories based on the principle of how or the process by which digital materials came into existence or rather were created [8]. One is digital material obtained through the process of digitization of physical material and the other is born digital. The terms digital preservation and digital archiving are not synonymous and cannot be used interchangeably. There is a fundamental difference between the two. As defined by Digital Preservation Coalition, digital preservation refers to a series of adopted management activities necessary to ensure a continued access to digital materials for as long as necessary, while digital archiving is referred to as the process of creating backup as an ongoing maintenance as opposed to strategies for long-term digital preservation.

Digital materials comprise "texts, databases, still and moving images, audio, graphics, software and web pages, among a wide and growing range of formats" [9]. Most of these materials are "frequently ephemeral, and require purposeful production, maintenance and management to be retained". It has been acknowledged that such "resources have lasting value and significance, and therefore constitute heritage that should be protected and preserved for current and future generations" [10]. Even though there lies a very thin line of distinction between "born digital" and "made digital" publications, this distinction is upheld here to illustrate the different problems inherent with each type of document having their own complexities that require different approaches for preservation. It needs to be highlighted that the "born digital" and "made digital" resources need different preservation methods. On one hand, the "born digital" resources do not have any print backups and it may only work with a specific type of hardware, software. On the other hand, the "made digital" or "digitized" resources are usually converted from print documents and we need to capture the authenticity of the original content while retaining the original source for future innovations.

Digital publishing has created new and different categories of document types from simple textual documents to multimedia document types, with varying types of media embedded therein. The advent of new and exciting digital formats such as graphs, audio, video has transformed the document types into very colorful, yet complex multimedia documents at times. The embedding of audio, video, graphics, sound and other aids have enhanced the quality of digital documents. Digital publishing encourages these types of new innovations. Digital preservation community is, therefore, forced to strategize ways of preserving these unique document types. For example, decision is to be made regarding aspect(s) of the metadata unfolding the digital materials and its preservation format. There are two major aspects to the preservation of digital materials: the practical and technological challenges that are within the control of the owning organization, and the political and organizational context that is not. Because of the seeming need to understand computer processing technology in depth for digital preservation, many conservators might think it outside their reach [11].

CHALLENGES OF DIGITAL PRESERVATION

The task of preserving digital assets adds a new set of challenges for libraries and archives to the existing

task of preserving a legacy of materials in traditional formats. Unfortunately, the needs and constraints of repositories shape the digital preservation strategies with little consideration for the requirements of current and future users of digital scholarly resources. A lot of challenges are faced in preservation of digital information. Some of the major challenges are as bellow.

TECHNOLOGICAL CHALLENGES

The large amount of electronic records demands new mass storage media. The problem is finite durability of media and that there is also often a lack of knowledge about the durability of different media for storage. The software problem concerns the rapid development of new versions of the software product and occasionally changeovers to new software products. Connected to the problem with rapidly changing software, there is the problem with changing formats, which in the long run can lead to the consequence that archival information becomes unreadable [12].

LEGAL CHALLENGES

An important requirement on digital records is evidence. How to ensure authenticity and how to preserve evidence is emphasized by many authors. The rapidity and the ease with which changes can be made mean that there are challenges associated with ensuring the continued integrity, authenticity, and provenance of digital materials [13].

ORGANIZATIONAL PROBLEMS

All of the challenges associated with digital preservation are not technical. A majority of them are organizational in nature. The first line of defense against loss of valuable digital information rests with the creators, providers, and owners of digital information. There is also a matter of inter-organizational problems to cope with. For example, government organizations want to offer electronic services to the citizens that require collaborations as

the services assume information from more than one organization [14].

CONTEXT AND METADATA

It should be possible for a user in the future to understand the context of the information and the processes that created it. How can we determine what amount of evidential historicity is required and how to capture and retain it? We do it with the help of storing contextual metadata along with the digital information.

MIGRATION

Hardware and software may become obsolete or non-maintainable in future. There may also be a preference for a better hardware/software. In such a scenario, we will have to migrate the digital asset to the new system. It is another biggest challenge in digital preservation. Digital migration is broadly described as moving into the new or advanced digital standard without compromising the realities of the current standard. It may be described as "a set of organized tasks designed to achieve the periodic transfer of digital materials from one hardware/software configuration to another or from one generation of computer technology to a subsequent generation". Migration is the most commonly used preservation strategy, especially for non-interactive digital objects such as images, databases or textual documents.

ESTABLISHED STANDARDS

Absence of established standards, protocols, and proven methods for preserving digital information is another challenge. Digital library research has focused on architectures and systems for information organization and retrieval, presentation and visualization, and administration of intellectual property rights. The critical role of digital libraries and archives in ensuring the future accessibility of information with enduring value has taken a back seat to enhancing access to current and actively

used materials. As a consequence, digital preservation remains largely experimental and replete with the risks associated with untested methods; and digital preservation requirements have not been factored into the architecture, resource allocation, or planning for digital libraries.

SKILLS AND EDUCATION

The non-availability of required skills and resources in some parts of the world has the effect that curators of information do not receive sufficient training in electronic systems. The level of education of the general population in many developing countries is also a problem. This creates a challenge for digital preservation both in terms of collection building, especially for end-user submissions, and dissemination. The skill of higher levels for managing the computer infrastructure is also essentially required. Novel solutions are needed for both these problems to make digital archives effective.

COST

The cost implication has an important bearing on digital preservation because no institution has access to unlimited funds. Cost implications depend mainly on the socio-economic dimension of the information along with the storage, hardware and software cost; and cost of the human resources. The ability to employ and develop appropriate skills in digital preservation entails training and re-training of information professionals as the technology changes.

CONCLUSIONS

There can be a critical and cumulative weakness in our information infrastructure despite huge information technology investments. Long term preservation of digital assets is limited by short media life, obsolete hardware and software, slow read times of old media, and defunct web sites. Indeed, the majority of products and services on the market today did not exist five years ago. More

importantly, we lack proven methods to ensure that the information will continue to exist, that we will be able to access this information using the available technology tools, or that any accessible information is authentic and reliable.

A lot of problems remain to be solved before secure and efficient digital preservation will be brought about. It seems that although much of the challenges associated with digital preservation is strategic, organizational, and structural and not only technical, the published research concentrates at solving the technical issues. Digital preservation requires, in addition to the technological development, elaborated strategies, new workflows and organizational structures, standards and common metadata, new specific competences, and close co-operation across different professions from traditional preservation management to computing science.

REFERENCES

1. Hedstrom, M. "Digital Preservation: Problems and Projects" paper presented at DL20, Tsukuba, Japan, March 2001, www.dl.ulis.ac.jp/DLjournal/No_20/1-hedstrom/1-hedstrom.html

2. Kerry, Angela. "Digital Preservation : The Research and Development Agenda in Australia" http://researchbank.swinburne.edu.au/vital/access/services/Download/swin:2873/DS2

3. http://osarome.blogspot.in/2011/10/challenges-of-digital-preservation-in.html

4. IDC White Paper "The Diverse and Exploding Digital Universe" sponsored by EMC, March 2008 http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf

5. New International Webster's Comprehensive Dictionary of the English

Language. (1999). Delux Encyclopedic edition Florida: Trident Press, 1999.

6. Digital Preservation Briefing Paper, 2006. Retrieved December 12, 2008 from http://www.jisc.ac.uk/publications /publications/pub_digipreservationbp.aspx

7. Smith, B. (2002). Preserving Tomorrow's Memory: Preserving Digital Content for Future Generations. Information Services CMG1Use, 22, pp. 133-139.

8. Digital Preservation Coalition (DPC) . "DPC history" 2002. available at www.dpconline.org/graphics/about/history. html.

9. UNESCO. Vol. 27, "Guidelines for the preservation of digital heritage", 2003. available at: http://unesdoc.unesco.org/ images/0013/001300/130071e.pdf (accessed 27 September 2005).

10. UNESCO (2003b), "UNESCO Charter on the preservation of the digital heritage: Document 32 C/28", available

at:http://portal.unesco.org/en/ev.php (accessed 6 October2005).

11. Keene, S., 2001. Preserving digital materials: confronting tomorrow's problems today. The Conservator, no. 26, pp. 93-99 http://www.suzannekeene.info/articles/digi pres.pdf

12. Gibbs, R. & Heazlewood, J. "Electronic records: Problem solved?" A report on the public office Victoria's electronic records strategy. The Journal of the Australian Society of Archivists, 27.1 (1999) 1-13.

13. Beagrie, N. A continuing access and digital preservation strategy for the Joint Information Systems Committee (JISC) 2002-2005. Retrieved November 22, 2005, from http://www.jisc.ac.uk/index.cfm?name=stra tegy_preservation

14. Asproth, V. (2006a). "Inter-organizational management and decision-making." Systemist,   28.2(2006) 4-12.