

Trends in Preserving Scholarly Electronic Journals

1. Golnessa GALYANI MOGHADDAM

Shahed University

Dept. of Library and Information Science, Shahed University, Persian
Gulf Highway, Tehran, IRAN, Postal code: 3319118651

E-mail: g_galyani@yahoo.com

2. Mostafa MOBALLEGHI

Karaj Islamic Azad University

Dept. of Industrial Management, Islamic Azad University (IAU) - Karaj
Branch, P.O. Box 31485-313, Karaj, IRAN.

E-mail: m_moballeghi@yahoo.com

Summary:

Scholarly electronic journals have become the largest and fastest growing segment of digital collections for most libraries. Many issues and concerns for managing electronic journals relate to preserving and providing continued access to them. The preserving of scholarly electronic journals is a complex issue with various aspects and is largely different from archiving of print-based scholarly journals. In this paper the author deals with issues concerning archiving of scholarly electronic journals. The purpose of this paper is to identify and discuss different issues related to preserving scholarly electronic journals. The following issues are discussed: differences between print and digital media, shift in the responsibility of archiving, copyright and intellectual property rights, cost of archiving, expertise, selection, redundancy, organizational issues, etc. Technical issues and challenges related to digital preservation include a lack of practical implementations of preservation standards and a lack of technical knowledge, in general, of what information is required to support the digital preservation process within organizations. Nevertheless, digital preservation has received considerably more prominence in recent years, gaining the attention of entities such as national libraries, national archives and other organizations.

Keywords: Digital Preservation, Electronic Archiving, Scholarly Electronic journals.

Introduction

The preserving of scholarly electronic journals is a new ground of research with various aspects. The purpose of preserving the electronic journals is to ensure that they remain accessible now and future. In this paper the author addresses

some of the important issues surrounding preservation of digital resources especially scholarly electronic journals. The author deals with issues concerning archiving of scholarly electronic journals such as differences between print and digital media, shift in the responsibility of archiving, copyright and intellectual property rights, cost of archiving, expertise, selection, redundancy, organizational issues, etc. Following the discussion of challenges and issues by details, the author attempts to trace the trends in digital preservation of scholarly electronic journals.

Background

The word *archiving* often refers to the process of storing physical objects, generally though not exclusively paper-based, in a physical location, such as a room or a building, to maintain that object's physical integrity and its intellectual context as could be represented by other objects within the archive¹. Digital archiving has little to do with physical objects or physical storage and it is different from the traditional meaning of *archive*, even some experts prefer to use *digital preservation* instead of *digital archiving*. The vocabulary such as *digital preservation* are being used in handbook of Digital Preservation Coalition² to define all the activities employed to ensure continued access to digital resources which have retained properties of authenticity, integrity and functionality. This is a richer interpretation and means that there will need to be more thought and preparation given to what resources are stored, how they are maintained and subsequently accessed and by whom.

In a new definition by S. Rabinovici-Cohen and his colleagues digital preservation comprises two aspects: *bit preservation*, which is the ability to access the bits of the digital record, and *logical preservation*, which is the ability to use and understand the data in the future. In addition, logical preservation must support tracking the provenance of a record and ensuring its authenticity and integrity. Bit preservation issues are mostly well understood and are supported by existing products used to migrate data between different generations of storage technologies. In contrast, logical preservation is still mostly an unsolved problem³.

¹ Seadle, 2006.

² The Digital Preservation Coalition (DPC) was formed in July 2001 to raise awareness of the issues raised by the need to keep and to re-use for a decade or more digital assets and resources which institutions have created or purchased. Further information on the DPC is available from its website (<http://www.dpconline.org/>).

³ Rabinovici-Cohen, et al, 2008.

With a view into technological context today, developments in information technology have obviously changed the traditional system of publishing, distributing, and even the use of scholarly journals. The initial communication for publishing a paper is so quick now, especially with the help of e-mail. An accepted manuscript can be accessed online before the date of publication. Even the patterns of use of scholarly journals are changing in the digital environment⁴. With the impact of information technology preservation of scholarly journals is more complicated than print-based journal and it has social, economics, legal, organizational and technical dimensions.

In January 2008, *Portico* and *Ithaka* invited 1,371 library directors of four-year colleges and universities in the United States to respond to a survey examining current perspectives on preservation of e-journals. A strong response has yielded interesting findings that we now share with the community in the hope they will spark useful discussion among library directors, funders, and administrators regarding strategic library priorities. The survey finds that a large majority of library directors across the spectrum strongly agree or agree that the potential loss of e-journals is unacceptable, and a significant majority believe their own institution has a responsibility to take action to prevent an intolerable loss of the scholarly record. Most larger libraries responding now support one or more e-journal preservation initiatives; however, the majority of respondents from smaller libraries have yet to support any preservation effort and secure permanent access to e-journals for their institutions. The survey shows that this majority is significantly uncertain about their options for e-journal preservation and how urgent is the need to act⁵.

Differences between Print and Digital Media

Electronic journals have many advantages over print journals. Online access allows for easier searching and retrieval of a topic. Electronic journals can be accessed anywhere (given proper equipment and software), and they have the ability to link to other people and resources beyond locations or *place*. With the greater capability of the digital medium, however, the content of digital files may be lost to future scholars not just because the physical item deteriorates, but because the information cannot be extracted and interpreted correctly. A scholarly journal on the printed page can be viewed and read without any special equipment as long as one knows the language in which it is written. Digital scholarly journals, however, cannot be viewed without special equipment, such as a computer, an Internet connection, and the required software.

⁴ Liu, 2005; Tenopir, 2005.

⁵ Digital Preservation of E-Journals in 2008, 2008.

With the machine dependency, archiving of electronic journals is more complicated than archiving print journals. The life expectancy of digital media is another issue of concern. The short lifecycle of digital media is a threat for digital archiving because digital media become obsolete much faster than print media. The format of the digital resources can be damaged or lost and may no longer be intact, retrievable, understandable, or displayable. The technology used to store the publication is likely to become obsolete even before that happens⁶. Therefore, continued access to archived resources is a big issue in digital archiving, while *access* was not a big issue to traditional archiving.

Shift in the Responsibility of Archiving

Information technology has caused substantial changes in the traditional roles of libraries and publishers⁷. One of the major changes is a shift in responsibility of archiving from libraries to publishers in an electronic environment. Historically, archiving records and documents has long been the responsibility of librarians, and publishers largely shade away from this role. Several libraries hold many research journals in print from their first volumes. Few publishers have complete journal collections archived for posterity. In the electronic environment, publishers and producers of scholarly journals are practically undertaking the responsibility of archiving, however. Magie Jones (2003) has pointed out this issue as follow:

“the transition from purchasing print journals, which the library then owned forever, to licensing access to e-journals for a defined period of time has major Implications for libraries and publishers. In terms of archiving responsibilities, there are no longer any clear-cut distinctions between who should be doing what. There is a lack of clarity regarding responsibilities and uncertainty about precisely what libraries are paying for when they license journals. This has meant that the transition from print to electronic has been more problematic than it might otherwise have been.”⁸

With a historic view, the trends in responsibility of digital archiving have been as follow: (1) The people with long traditions of preserving physical artifacts (e.g. archivists, librarians, museum curators) increasingly recognized that it is their responsibility which is now digital. (2) The people with long traditions of managing computer-dependent data sets (e.g. scientific data center personal, technology managers) increasingly recognized that it is their responsibility.

⁶ Steenbakkens, 2005.

⁷ Steenbakkens, 2005.

⁸ Jones, 2003.

There is a debate over responsibility of digital archiving among all stakeholders, but at the case of electronic journal it seems that the publishers practically have to accept the responsibility of digital preservation; as in digital environment, electronic publications (particularly electronic journals) are not physically owned by libraries. Although libraries traditionally owned the resources forever once they paid to publisher, now they license access from the publisher. In fact, licenses are an agreement for legal use of electronic resources not for ownership and publishers will remain the owner of electronic resources. Libraries as subscribers are therefore concerned that publishers do not consider the archiving and preservation of these works and include archiving and perpetual access to back issues in licensing of these works.

Copyright and Intellectual Property Rights

Copyright and other intellectual property rights (IPR) are two important issues because of their substantial impact on digital preservation. We know that copyright law was originated long before there was a thought of the World Wide Web. Copyright seems to be established well for traditional archiving but not for electronic materials. The copyright and intellectual property rights issues for digital materials are much more complex and significant than for traditional media. If these issues are not addressed adequately, preservation will be curtailed. Both contents of digital resources and their associated software needs to be taken into consideration. It may be noted that the current archiving initiatives (such as JSTOR, Portico, E-Print Repositories, LOCKSS, OCLC Digital Archive, JISC, PubMed Central, Open Access Model, e-Depot, etc), have adopted many divergent approaches to preserving intellectual contents over time because of complexity of copyright law in digital environment⁹.

Copyrights issues have not got a quick solution in digital preservation, as copyright law allows only fair use and it can prohibit a successful preservation to some extents. Some experts suggest to put away copyright in digital preservation or make some changes in law, though it is not easy to do. They reason if current law does not allow copying for digital preservation, the most obvious solution is to change the law and if libraries want to preserve information, they need to be able to carry out the required activities¹⁰.

Although making changes in law or licensing practice is difficult, rights holders and libraries have to understand and cooperate with each other to progress. There are many stakeholders who may have an interest in archiving electronic journals. Mary Feeney describes in detail the stakeholders as authors, publishers, libraries, archive centers, distributors, networked information service providers, IT suppliers, legal depositories, consortia, universities, and

⁹ Galyani M., 2008.

¹⁰ Muir, 2004.

research funders. Feeney also suggested that the relationship of the stakeholder to the digital material archiving needs to be taken into consideration¹¹.

Cost of Archiving

The other important issue in digital preservation is cost of archiving. Digital preservation is essentially about preserving access over time and therefore the costs for all parts of the digital life cycle are relevant. Of course, digital access has many advantages over paper-based or microform access in terms of convenience and functionality, however, providing continued access is an important concern for digital librarians. Cost of digital preservation seems to be much higher than the cost of traditional preservation. Access to digital resource with the rapid technological changes is not easy and needs expert staff and considerable expenditure on technological needs.

Mary Feeney (1999) gives a thorough breakdown of cost considerations based on one of the studies commissioned by the Digital Archiving Working Group (DAWG). She pointed out:

“One clear message that has emerged is that a great deal of money can be wasted if digitization projects are undertaken without due regard to long-term preservation. It is now relatively easy to produce digital versions of texts or images. However, if there is no plan in place for archiving the digital files, long-term preservation will be expensive, or may even result in the work having to be repeated”.¹²

Calculation of costs in digital archiving is not easy, however, is a valuable and necessary task to establish a cost-effective and reliable business model. Costs for *maintaining* the digital copy also need to be considered from the beginning whether those materials are produced as a result of digitising analogue materials or they are *born digital*. It may be noted that other issues such as organizational mission and goals including the type and size of collections, the level of preservation committed to and the quantity and level of access required, and time frame proposed for action should take into consideration.

Expertise

Digital preservation needs high skilled staff while in the traditional archiving the scenario was different. Montgomery and Hedstrom pointed out

¹¹ Feeney, 1999.

¹² Feeney, 1999.

“The need for digital preservation expertise is high: asked to rate staff as expert, intermediate, or novice, only 8 of the 54 institutions considered their staff at the expert level”¹³.

It is obvious that the ability to employ and develop staff with appropriate skills is made more difficult by the speed of technological change and the range of skills needed. Continuous training and *learning by doing* are the methods that can be adopted while both methods have their own limitations. Libraries need to ensure their existing staff and members can develop, and continue to develop, the range of competencies they need to manage the digital materials in their care.

Selection

Selection is another important issue in electronic archiving. The huge quantity of information being produced digitally, its variable quality, and the resource constraints on those taking responsibility to preserve long-term access make selectivity inevitable for archiving. Traditionally, lack of selection for preservation may not necessarily mean that the item will be lost, but in the digital environment non-selection for preservation will almost certainly mean loss of the item. Although not all resources can or need to be preserved forever, some will not need to be preserved at all, others will need to be preserved only for a defined period of time, and a relatively small sub-set will need to be preserved indefinitely. The decision should be made as early as possible to help save resources for the most valuable digital assets.

In digital preservation where there are multiple versions of an item, decisions must be made in selecting which version is the best one for preservation, or whether more than one should be selected. The importance of selection has been acknowledged by many stakeholders, e. g. the National Library of Canada (NLC)’s guidelines state,

“The main difficulty in extending legal deposit to network publishing is that legal deposit is a relatively indiscriminate acquisition mechanism that aims at comprehensiveness. In the network environment, any individual with access to the Internet can be a publisher, and the network publishing process does not always provide the initial screening and selection at the manuscript stage on which libraries have traditionally relied in the print environment. Selection policies are, therefore, needed to ensure the collection of publications of lasting cultural and research value”¹⁴.

Redundancy

¹³ Hedstrom and Montgomery, 1998.

¹⁴ NLC 1998.

In traditional archiving, some level of redundancy with multiple copies was inevitable in different repositories, but the story is different in the electronic environment. Some authors, such as Dale Flecker (2001), believed that there was large-scale redundancy in the storage of journals in the print era, as many different institutions collected the same titles. Theoretically, in a digital environment, a single institution can provide worldwide access and accept preservation responsibility, although there is a debate whether a level of redundancy should exist in the digital environment¹⁵. In order to avoid the danger of losing access over time, at least one copy of materials should be stored in two different repositories. Librarians should make clear who will undertake preservation responsibility and for what period of time. Making appropriate documentation for each level of preservation, selection process and responsibility can give some assurance to have successful preservation strategies.

Organizational Issues

There are many organizational issues regarding digital preservation. Digital preservation requires new workflows, new skills and close co-operation across different professions ranging from traditional preservation management skills to computing science. The organizational structure to support this is not yet in place¹⁶. There is lack of clarity in roles and responsibilities between organizations and between different stakeholders. The Digital Preservation Coalition (DPC) carried out a UK-wide survey to assess the nation's preservation needs in 2006. One striking result of the survey is the common lack of clarity in responsibilities for digital preservation, which has been seen by a majority of the respondents as a barrier to digital preservation¹⁷.

We may be noted that although the situation in digital archiving has been improved since 2006, the organizational issues still need to be taken into consideration. Organizations need to understand digital preservation needs, expertise, technological infrastructures, costs and prepare proper strategies to ensure a successful digital preservations.

Discussion and Conclusion

The preserving of scholarly electronic journals is a complex issue with various aspects and is largely different from archiving of print-based scholarly journals. With a broad view, preserving of scholarly journals has social, economics,

¹⁵ Flecker, 2001.

¹⁶ Hockx-Yu, 2006.

¹⁷ Waller and Sharpe, 2006.

legal, organizational and technical dimensions. The issue of differences between print and digital media, shift in the responsibility of archiving, copyright and intellectual property rights, cost of archiving, expertise, selection, redundancy, organizational issues are discussed and covered in this paper. Digital preservation seems to be a complex process and there are many unsolved organizational, managerial and technical issues that make digital preservation a challenging task for all stakeholders.

Technical issues and challenges related to digital preservation include a lack of practical implementations of preservation standards and a lack of technical knowledge, in general, of what information is required to support the digital preservation process within organizations. The challenges associated with digital preservation are not purely technical. In order for digital archives to be sustainable over time, the organizations responsible for the archives must have appropriate expertise, resources, and political/institutional mandate to carry out the work required. Given the cost and complexity of digital archives, as well the potential to exploit the rich sets of relationship across individual collections, coordination of work across social boundaries (institutional, regional, disciplinary, organizational and professional) is also important.

We may note there are some threats for long-term availability of electronic resources. Data mismanagement, technological dependency, media degradation and technological obsolescence have all threatened the long-term accessibility of resources stored in digital formats.

Nevertheless, digital preservation has received considerably more prominence in recent years, gaining the attention of entities such as national libraries, national archives and other organizations. It has to come to be recognized as a legitimate and essential area of research and development. Many stakeholders of scientific publishing have begun to consider importance of electronic archiving and take initial steps to meet their responsibility effectively. The new concerns of electronic archiving led to a series of meetings over the past few years among publishers, librarians, and technologists sponsored by a variety of organizations. In order to manage the archiving issues, different initiatives and projects (such as JSTOR, Portico, E-Print Repositories, LOCKSS, OCLC Digital Archive, JISC, PubMed Central, Open Access Model, e-Depot, etc) were created by various organizations and institutions.

Finally, digital preservation requires new workflows, new skills and close cooperation across different professions ranging from traditional preservation management skills to computing science. There is a need toward more awareness of digital preservation among all stakeholders. This field still is in its infancy.

References

- Feeney, M. (1999), "Towards a National Strategy for Archiving Digital Materials", *Alexandria*, Vol.11, No.2, pp. 107–122.
- Flecker, D. (2001), "Preserving Scholarly E-Journals", *D-Lib Magazine* (September) Vol.7, No.9. Available at: <http://www.dlib.org/dlib/september01/flecker/09flecker.html> (accessed September 20, 2008).
- Digital Preservation of E-Journals in 2008: Urgent Action Revisited Released (2008), Available at: <http://digital-scholarship.org/digitalkoans/2008/06/06/> (Accessed July 17, 2009).
- Galyani Moghaddam, G. (2008), "Preserving Scientific Electronic Journals: A Study of Archiving Initiatives", *The Electronic Library*, Vol.26, No.1, pp.83-96.
- Jones, M. (2003), "Archiving E-Journals Consultancy - Final Report Commissioned by the Joint Information Systems Committee (JISC)", October. Available at: http://www.jisc.ac.uk/uploaded_documents/ejournalsfinal.pdf. (Accessed August 18, 2008).
- Hedstrom, M. and S. Montgomery (1998), "Digital Preservation Needs and Requirements in RLG Member Institutions" A study commissioned by the Research Libraries Group. December 1998. Available at: <http://www.rlg.org/preserv/digpres.html> (Accessed August 17, 2008).
- Hockx-Yu, H. (2006), "Digital Preservation in the Context of Institutional Repositories", *Program: Electronic Library and Information Systems*, Vol. 40, No. 3: pp. 232-243.
- Liu, Z. (2005), "Reading Behavior in the Digital Environment: Changes in Reading Behavior Over the Past Ten Years", *Journal of Documentation*, Vol.61, No. 6, pp.700–712.
- Muir, A. (2004), "Digital Preservation: Awareness, Responsibility and Rights", *Journal of Information Science*, Vol.30, No.1, pp. 73-92.
- NLC (1998), National Library of Canada, Electronic Collections Coordinating Group. *Networked Electronic Publications Policy and Guidelines*, October 1998. Available at: <http://www.nlc-bnc.ca/pubs/irm/enepg.htm> (Accessed July 17, 2009).
- Seadle, M. (2006), "A Social Model for Archiving Digital Serials: LOCKSS", *Serials Review*, Vol.32, No.2, pp.73-77.
- Steenbakkens, J.F. (2005), "Digital Archiving in the Twenty-First Century: Practice at the National Library of the Netherlands," *Library Trends*, Summer.
- Tenopir, C. (2005), "Inundated with Data", September. Available at: <http://www.libraryjournal.com> (Accessed December 15, 2008).
- Rabinovici-Cohen, S. et al. (2008) "Preservation DataStores: New storage paradigm for preservation environments" *International Business Machines Corporation (IBM), J. RES. & DEV.* Vol. 52 No. 4/5 JULY/SEPTEMBER.
- Waller, M. and Sharpe, R. (2006), "Mind the Gap: Assessing Digital Preservation Needs in the UK, Digital Preservation Coalition, York", available at: www.dpconline.org/docs/reports/uknamindthegap.pdf (accessed January 16, 2008).