Annals of Library and Information Studies
Vol. 59, September 2012, pp. 204-211

Role of consortia in preservation of e-journals

Ramesh C Gaur[a] and ManoramaTripathi[b]

[a]University Librarian, Jawaharlal Nehru University, New Meharuli Road, New Delhi – 110067,

Email: rcgaur@mail.jnu.ac.in ;rcgaur66@gmail.com

[b]Deputy Librarian, Jawaharlal Nehru University, New Meharuli Road, New Delhi – 110067,

Email: manoramatripathi@mail.jnu.ac.in

The present paper briefly describes the importance of library consortia and enlists main consortia of the country. It mainly dwells upon the various issues related to the preservation of digital content of e-journals. It suggests that various Indian library consortia in the country should take lead and initiate action for preserving e- journals for future.

Keywords: e- journals, consortia, archiving; digital preservation

Introduction

The origin of library consortia can be traced back to 1886 when Melvil Dewey dwelt upon "library cooperation" in *Library Journal* and so did E.A Mac in 1887 in his article "cooperation versus competition in *Library Journal*[1]. R.B. Downs expressed his views on library cooperation in his paper "one for all: a historical sketch of library cooperation 1930- 1970" which was included in the 1939 symposium organized by ALA-The library of tomorrow. Libraries, worldwide, participated in some formal or informal interlending programmes long before the onset of Internet technology; this was popularly known as library cooperation or resource sharing. These resorted to reciprocal borrowing and use of material from one another in order to meet the requirements of documents which were not available with individual libraries. Academic libraries have always formed consortia for the purpose of sharing existing physical resources like books, journals etc., held by member libraries. The formation of any consortia is based on the fact that a group of libraries has a combined set of resources which is greater than the resources of any single member library[2].

Cooperative acquisition programme of NISCAIR erstwhile INSDOC, establishment of DELNET in 1988 and INFLIBNET 1991 besides establishment of a number of local libraries networks such as CALIBNET (Calcutta), BONET (Bombay), PUNENET (Pune), MALIBNET (Madras), MYLIBNET (Mysore), HYLIBNET (Hyderabad), ADINET (Ahmedabad), and countrywide ones like ERNET (Educational and Research Institutions), SIRNET (CSIR Laboratories) are some of the early library cooperation initiatives in India.

With the advent of Internet, the academic libraries have started forming consortia to provide common access to electronic resources across the Internet. The consortia arrangements enable libraries save money, survive financial constraints, and eliminate redundant resources and activities. "The primary purpose of establishing a library consortium is to share physical resources including books and periodicals amongst members. However the mode of cooperation has gone under transformation with the infusion of new information technology from print based environment to digital environment. The emergence of Internet, particularly the World Wide Web as a new media of information delivery (has) triggered proliferation of web based full text online resources."[3]

The costs of the e-resources were quite high and it was beyond the reach of almost all libraries to subscribe required e-resources within available budgets. This was another challenge for Indian libraries that were already facing the problem of increasing cost of the publications and shrinking budget. This lead to find out alternative models, hence library consortia with the support of state funded agencies emerged in India.

Importance and use

Library consortia reflect shift from organizational self-sufficiency to "collaborative survival mode" among the libraries[4].

The advantages of library consortia are as under:

- Promote collaboration in different library activities.
- Better sharing of existing resources and jointly acquiring new resources at great savings.
- Negotiating and finalizing terms and conditions with the publishers.
- Streamlining and consolidating technical services through compilation of union catalogue.
- Professional development and training.
- Promotion of new standards for managing usage statistics
- Undertaking digitization projects, setting standards for best practices and providing training in them[5].

Library consortia in India

The International Coalition of Library Consortia (ICOLC) which is an informal group at present lists approximately 200 library consortia from around the world. It supports participating consortia by facilitating discussion on issues of common interest. (http://icolc.net/).After China, India is the second largest subscriber of Consortia based library subscription of e-resources. Some of the important library consortia in India are as follows:

- INDEST-AICTE Consortium
- CSIR E-Journals Consortium
- UGC Infonet Consortium
- FORSA Consortium
- IIM LibrariesConsortium
- Medical Consortium (NML)
- ICAR- CeRA

The further details of some of the major Indian library consortia are given in Table 1.

Threats to e-journals' content

Technological obsolescence

The physical media of digital information is vulnerable to fast damage; the format of e- journals may get lost forever. Even the technology which is used to store e- journals may become obsolete. The functionality which is required to access, browse and use information of e-journals may get lost. If the functionality of specific hardware and software is not available, the content cannot be used[6].

Print vs. online

In good old days libraries used to subscribe print journals. Bound volumes of these journals are still available with many libraries in India. E- Journals are convenient for the users to browse and access economical in terms of management cost and storage space. But there are many problems and challenges associated with e-journals. The digital information which is contained in electronic journals is very fragile. JaJa and Song (2009) have clearly stated that digital information is very fragile and can be lost easily due to hardware and software obsolescence and frequent upgrading[7]. The vagaries of nature-fires, floods and hurricanes can also lead to loss of data.

In print environment, there is no distinction between preservation of journals for present or future use. There is lot of replication and redundancy in journal collection as many libraries subscribe and purchase the same journal titles. The copies of journals being saved for future generations are the same copies which are being used and read by the current generation of users. In print, most of the activities like binding, repair, handling and shelving which are routinely followed for maintaining journals and serving current generation of users also ensure their preservation and availability for use by future generations for next 100 years or so. In online environment, hardly any library actually holds digital copies locally.

Perpetual access vs. ownership

Libraries just get access; they do not own the content, as they do in case of print subscriptions. There is no local hosting of e-journals and databases. These are accessible from the publishers' websites, most of which are located outside the country. These websites provide search and browse interface to accesses the journals. Libraries have absolutely no control over the content they subscribe or license. After spending huge amounts of money on annual subscriptions, libraries just get access to content not ownership of content.

Access is provided for a fixed period of time, that too, according to terms and conditions specified by the publishers. The terms and conditions specify method of access-single user or unlimited number of concurrent users; define authorized users and services like downloading, printing, interlibrary loan and document delivery. While negotiating, the libraries pay least attention to archiving and long term preservation plans of the publishers. Libraries are more bothered about getting access to back issues; access to unlimited number of concurrent users, clauses of downloads and document delivery services,

Table 1—E-resources Consortia in India

Sl. No.	Name	Year of Inception	Subject/Coverage	No of journals/ publishers covered	No of Members	Functions
1.	INDEST http://paniit.iitd.ac.in/indest/	2003	Science and Technology, Management	12,000+ journals from 25+publishers; 6 bibliographic and citation databases	1363	e-content licensing, Interlibrary loan, training, cataloguing services
2.	UGC-INFONET Digital Library consortium http://www.inflibnet.ac.in/ econ/index.php	2003	All subject areas	7000+journals from 26 publishers and aggregators ; 10 bibliographic/citation databases	321 universities/ colleges	E- journal licensing Interlibrary loan, training, cataloguing services
3.	ERMED(Electronic Resources in Medicine) Consortium http://www.nmlermed.in/	2008	Medical Sciences and allied fields	1812 journalsfrom 9 major publishers	98 medical colleges/ institutes	E- journal licensing
4.	National Knowledge Resource consortium http://nkrc.niscair.res.in	2001	Science and Technology	10,000 online journals from 28 publishers	39 CSIR and 24 DST institutes.	E- journal licensing
5.	RGUHS-HELINET Consortium		Medical/Health Sciences	600+ journals from 11 publishers	276+ medical colleges	E- journal licensing
6.	CeRA http://cera.iari.res.in/	2007	Agricultural Sciences	2000+from 8 publishers.	126 institutions	E- journal licensing. Interlibrary loan, training, cataloguing services
7.	DELNET		Management, Pharmacy, engineering, science and Technology, Medical Sciences		100+	E- journal licensing Interlibrary loan, training, cataloguing services

usage statistics and discounts. Libraries hardly pay any attention to crucial issues of archiving e- journals' content. The agreements which are signed by libraries and publishers do not include any mention of archival policies and rights while many do not have a reference to perpetual access.

Libraries still focus on widening access to e-resources; they do acknowledge the importance of long term digital preservation but expect someone else i.e. national libraries, publishers or consortia to undertake digital preservation activities on behalf of them.

In 2011-12, JNU spent Rs. 4.00 crores approximately on subscription to e- resources. If one extends the expenditure trend beyond this university, the total expenditure which consortia and libraries across the higher education community in India make annually in licensing access to e-resources is astronomical and deserves to be noted.

The important difference between licensing access to an e-journal in contrast to purchasing and owning a print issue, is that until and unless there is some clear, unambiguous legal agreement between the libraries and publishers, the access to online content cannot be guaranteed forever[8]. Since the libraries do not have ownership of the online resource provided by the publisher, libraries have no right to archive it. Of course, in some cases, the publisher guarantee perpetual access of previous years, it is little ambiguous as it may imply perpetual access as long as the publisher's company exists. But as library professionals, we are concerned about getting the access perpetually-as long as the library exists. Even if the libraries get additional copies of online journals for archiving, they do not have the infrastructure, funds and technical expertise to undertake long term digital preservation procedures and activities.

Change of ownership or no ownership

In case a journal gets ceased or is acquired by another publisher or merged with another journal or the publisher gets liquated or some natural disaster strikes, the libraries will be at risk of losing content of online journals and databases which will eventually affect the researcher community of the country. Even if a publisher has an archiving program, it is not guaranteed that all its titles will be covered under the archiving program. Earlier the aggregators /publishers like Proquest, EBSCO, LISA used to provide CD-ROMs with annual subscription to the libraries; these CDs are no longer provided. Besides it is not at all prudent to completely rely on publishers for archiving and long term preservation of e- journals. Many publishers maintain archives of back files at network sites, but there is no guarantee that these files will be available permanently. Also, there exists the issue of technical expertise. The publishers' traditional expertise lies in creating and distributing information. Many publishers especially the small professional organization and scholarly societies may not have the plans, vision, infrastructure to create and maintain e- journals' archive and go for long term preservation. The cost of digital preservation is much higher than the preservation of printed material. Since technologies swiftly evolve, ensuring and maintaining access to digital resources is a daunting task. It entails additional expenditure, investment on expert staff, hardware, software and storage facility etc.

Paperless or e-only libraries

The researchers, teachers and students depend upon accumulated scholarship of the past in order to generate new knowledge. There is apprehension which is very much justified, across the academic world over the long term accessibility of e-resources. The academia is pensive about the possible gaps in the availability of e- content which may adversely affect their work.The libraries can't shift their localities to e-only forms—they continue to procure, process and store print journals; at the same time invest heavily on e- journals. This is certainly not a viable proposition. The libraries feel it is too early to rely exclusively on e-forms only. The "e" of e- journals stands for electronic; but there is a danger which looms large that "e" could come to indicate ephemeral if we don't take concrete step to preserve digital scholarship.

Some studies have highlighted that 13% of Internet sources which were cited in 3 reputed journals could not be retrieved from the original hyperlinks just after 27 months of publication. That is why issue of archiving for long term preservation has gained tremendous importance and needs to be attended too. The long term preservation of the e-journals is of greater importance to the libraries and research community than the publishers.

Digital preservation of e-resources: Some initiatives

Some of the projects which have been started at international level for preserving e- journals' content have been discussed below:

LOCKSS

This network is based on the concepts that replication of digital content at various locations can keep the content safe for future use. It uses open source software developed by Stanford University, which preserves digital content in a network of libraries just as libraries keep multiple copies of the same book. The software is free but membership to LOCKSS Alliance is essential for getting access to e-journals and certain other features. A library may use a general PC as LOCKSS Box to preserve digital content. A LOCKSS box performs the following functions:

- It collects content from the target websites by using a web crawler.
- It compares the content with the same content collected by other LOCKSS boxes.
- It acts as a web proxy or cache and provides access to the publisher's content.
- It provides a web based administrative interface which allows the library staff to target new journals for preservation, monitor the journals which are being preserved and control access to the preserved journals.

The publisher has to give permission for the LOCKSS system to collect and preserve the journal. This is done by adding a page to the journal's website, with a permission statement and links to the issues of the journal as they are published. The LOCKSS boxes have to know where to find this page and how far to follow the chains of web links. This is done via LOCKSS plugin that is supplied by the Stanford university LOCKSS team. http://www.lockss.org/about/how-it-works/Libraries preserve their digital content with LOCKSS Boxes in two different environments:

Global LOCKSS Network

Through this network the libraries preserve open access titles and those e- journals and e-books which they subscribe. At present, 9000+ e-journal titles from 510 publishers have been selected for preservation through Global LOCKSS Network. It is maintained by the Stanford University LOCKSS staff with financial support from the LOCKSS alliance. LOCKSS makes it economically and technically possible for even small libraries to preserve their own content. They need not outsource their own content to a third party service provider.

Private LOCKSS Networks

Libraries participate in PLN to preserve and archive their highly specialized collections of manuscripts, image collections etc. This kind of network offers libraries with synergistic collections, a mechanism to ensure the survival of their highly specialized content.

CLOCKSS archive

The CLOCKSS initiative was introduced in 2006 in order to offer digital preservation services to academic publishers. The goal is to preserve digital scholarship in cases where the journal is no longer available, due to merger, publisher has ceased operation or there is some major technical breakdown. The CLOCKSS archive provides this assurance via its secure network of content that can be accessed only when a trigger event is deemed to have occurred. It is because it makes all content triggered from the archive freely available to the world. It is a global archive that preserves content on behalf of all libraries and scholars worldwide. It preserves content in 12 chosen libraries across the globe to optimize the contents safety against political and environmental threats.

The various steps in CLOCKSS are as under:

- Publisher provides the CLOCKSS system access to either presentation or source. In order to allow, access, CLOCKSS crawls to access the presentation files-a CLOCKSS provided permission statement is added which informs what content is available for collection.
- Special CLOCKSS boxes located at RICE, Indiana, Stanford Universities ingest the content the publisher has made available.
- The content in each CLOCKSS BOX goes through a verification process to confirm that their versions of the content are identical to each other. This establishes the authoritative version of the content.
- Majority of the CLOCKSS boxes are preservation machines that perform the main storage and audit functions. After the quality of the content on the ingest machine is validated, it is collected from them by the preservation CLOCKSS boxes.
- The content is then preserved through a system of audit and repair. The CLOKSS boxes continually communicate over the Internet to audit the contents they are preserving. If the content in one CLOCKSS box is damaged or incomplete, the CLOCKSS box will receive repairs of the content based on other CLOCKSS boxes holdings and/or by referring to the publisher's presentation files.

When a trigger event occurs and the CLOCKSS Board decides to release the content from archive two things happen: content is automatically migrated to the newest format. Content is copied from the CLOCKSS boxes to a publicly available web server at a CLOCKSS host organization (currently the EDINA Data Centre, University of Edinburgh, and Stanford University).

At present the CLOCKSS archive has 9000 titles from 78 publishers. The number of items is 5 million. The content is captured in its original web published formats with all related files and links. In 2011, ASPET (American Society for Pharmacology and Experimental Therapeutics and the AMA (American Medical Association) asked the CLOCKSS board to trigger journals from their collections.

E-Depot system of National Library of Netherlands

The national library of Netherlands, the Koninklijke Bibliotheek (KB), located in The Hague, was founded in 1798.The library acts as a depository library for all print and digital publications. In 2002, Elsevier and KB signed an agreement for the archiving of all Elsevier e-journals. The main aims of the archiving agreement are as under:

- Bibliographic metadata about the publications may be included in the KB's online public catalogue and in the National Bibliography.
- Publications may only be used onsite at the KB and only by authorized users of KB.
- In the case of Open Access publishers and non-profit publishers, the onsite restriction does not apply.

- For authorized KB staff, both onsite and remote use is allowed. The archived digital publications may be used as a source for print or fax copies for interlibrary loan within Netherlands.
- Sending or transferring the electronic files outside the library by any other means is not allowed[6].

The agreement was later on signed with other publishers-Kluwer, Biomed Central, Blackwell Publishing, Taylor and Francis Group, Oxford University Press, Springer and Brill Academic Publishers.

E-Depot is the term for KB's infrastructure and organization for archiving digital publications including the workflow and procedures for handling and archiving digital content. The main components of e-Depot are DIAS and other subsystems which offer digital library functions. DIAS is Digital Information Archiving system software which has been developed by IBM for archiving digital content. DIAS is neither specific to KB nor even to libraries but rather a solution for digital archiving in general.IBM has been granted intellectual property rights for DIAS; and it serves as an incentive for IBM to brand, market, and update the archival system. The KB intends to develop e- Depot as "Trusted Digital Repository" (TDR).It is one whose mission is to provide reliable, long term access to managed digital resources to its designated community, now and in the future.http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf. The e–depot has archived more than 5000 e-journals,10 million articles; it has current storage capacity of 25 TB which will be extended to 1.5 PB in future http://www.kb.nl/dnp/edepot/operational/background/facts_and_figures-en.html

Portico

It is not for profit electronic service which was started by JSTOR in 2002.It focuses on preserving the intellectual content of electronic journals through file normalization and format migration (Fenton, 2006)[9].

In order to participate in Portico a publisher:

- has to sign a non-exclusive archiving license which gives Portico the right to ingest, normalize, archive and migrate the publisher's content,
- has to specify if Portico will serve as a perpetual access mechanism,
- has to supply source files in time,
- has to pay an annual fee.

Whereas a library has to:

- sign an archiving license agreement,
- has to pay an annual fee,
- has to provide details like IP ranges for user authentication.

At present, there are153 publishers on behalf of 2000+ societies and associations who have committed more than 13,919 e-journals and 129,890 e- books to the Portico archive; besides 750 libraries are also participative in Portico archive.http://www.portico.org/digital-preservation/the-archive-content-access/archive-facts-figures#page-1552

The participating libraries gain campus wide access when trigger events occur which lead to non-availability of journal titles from publishers' websites or any other source. The trigger events occur when a publisher ceases operations, or ceases to publish a title, or no longer offers back files of journal titles or suffers ccatastrophic and sustained failure of delivery mechanism. Some examples of triggered events are – Annals of Clinical Psychiatry, Australian and New Zealand Journal of Audiology. These are available to all participating libraries through the Portico website, regardless of whether the participating library had previously licensed the content or not.

Role of consortia in digital preservation of c-contents:

Consortia, libraries and the publishers are major stakeholders in the field of dissemination and distribution of scientific information to the user community at large. The consortia have an important role in ensuring perpetual access for future generations. Since they are entrusted with the responsibility of negotiating and finalizing subscription to online journals, they should take concrete steps to ensure perpetual access to future generations.

- As of now, consortia have not taken any decision or adopted any measure for preservation of e-journals' content. It is high time the library consortia in the country took concerted efforts to ensure perpetual access to the e- journals and online databases.
- The consortia should set up a digital preservation and archiving unit for e-journals where source or presentation files can be archived. They can lobby together and ask the publishers to cooperate and offer additional copies of digital content for

archiving for future use. The consortia, as it is obvious, have an edge over individual libraries. They should provide a safety net for all the online journals and databases which they subscribe. This clause can be incorporated in the license agreement which they sign with publishers.

- At present, some initiatives and projects have been started for preserving e-journals' content as described above, consortia should sign up and participatein a couple of archiving projects.
- They must seek information on publishers' policies with regard to the following issues:
  - o archiving policies for long term digital preservation and perpetual access,
  - o access post cancellation,
  - o e-journal transfer from one publisher to another,
  - o publisher gets liquidated;
  - o breakdown of publishers' servers due to natural disasters.
- The publishers like Elsevier, Oxford, Cambridge, Wiley, Taylor and Francis and Springer should be asked to host their content locally too in the country. They will not have any objection in it as some of them have licensed their content for local hosting in other countries. For instance, since 1996 M/s Elsevier have licensed their Science Direct services for local hosting in different countries; there are approximately 10 major library centres around the world which hold all Elsevier titles locally. Of course, these libraries are not official archives, their possession of the digital files provides considerable assurance that files will not disappear or be inaccessible to the academic fraternity.
- The consortia should take the initiative for sensitizing the policy makers and libraries towards the importance of archiving and long term preservation of e-journals. They should work to ensure that archiving and long term preservation are adopted in libraries and institutional strategies and policies for research and study.
- They should also train manpower for undertaking archiving and long term preservation activities; a small percent of annual budget should be earmarked for undertaking preservation activities.

The consortia should work to spread awareness among the libraries and parent institutions, identify and highlight the risk factors and if no timely action is taken, coordinate and work with publishers to create journals' archives and consider shared library solutions for archiving and access.

Conclusion

Concerted efforts are being made worldwide for preserving e-journals' content for posterity. As of now, no perfect e-journal archiving solution has emerged which can guarantee perpetual access to future generations. But as it is wisely said, a journey of thousand miles starts with a single step; library consortia in India should take lead and initiate measures for archiving and long term preservation of e journals' content in order to ensure that digital scholarship remains safe, accessible and usable for posterity. If no timely action is taken, access to back files of e-journals, scholarship, social and cultural heritage will be jeopardized. This will be totally against the 5th law of library science and will adversely affect all stakeholders-library professionals, users and funding bodies.

Time seems right for Indian Libraries to explore various options to create models for digital preservation of e-resources subscribed by them. Besides options like Portico, LOCKSS and CLOCKSS, the INDEST and INFLIBNET Consortium should also initiate work on setting up of some Digital Preservation Project for E-journals. To begin with they must insist various publishers to set up local servers to meet the needs of Indian Libraries[10].

References

1 Kopp J, Library consortia and information technology: the past, the present and the promise, *Information Technology and Libraries*, 17 (1) (1998) 7-12.

2 Potter W G, Recent trends in state-wide academic library consortia, *Library Trends*, 45 (3) (1997) 417–434.

3 Arora J and Agarwal P, Building digital libraries in a consortium mode: Towards a national consortium. In International Conference on Digital Libraries 2004: Knowledge creation, preservation, access, and management, New Delhi: The Energy Resources Institute, Vol. 1, pp. 292–311.

4 Allen B M and Hirshon A, Hanging together to avoid hanging separately: opportunities for academic libraries and consortia, *Information Technology and Libraries*, 17(1) (1998) 36-44.

5 Burke R, Library consortia and the future of academic libraries (2010) Available at: http://www.neal-schuman.com/academic/Burke2010.pdf (Accessed on 4 August 2012)

6 Steenbakkers J F, Digital archiving in the twenty-first century: practice at the National Library of Netherlands, *Library Trends*, 54(1) (2005) 33-56.

7 Jaja J and Song S, Robust tools and services for long term preservation of digital information, *Library Trends*, 57 (3) (2009) 580-94

8 JISC,e-Journals: Archiving and Preservation Briefing paper(2007) Available at: http://www.jisc.ac.uk/publications/briefingpapers/2007/pub_ejournalspreservationbp.aspx (Accessed on 14 August 2012)

9 Fenton E, Portico: An Electronic Archiving Service , Portico iPRES, October 10, 2006 http://ecommons.cornell.edu/bitstream/1813/3684/1/Eileen_Fenton-Portico.pdf (Accessed on 30th September 2012)

10 Gaur R C and Tripathi M, Digital preservation of electronic resources, *DESIDOC Journal of Library & Information Technology*, Special issue on Digital Preservation, 32 (4) (2012)