

## Are Research Libraries in India prepared in Digital Age?

Swapan Deoghuria

Computer Centre

Indian Association for the Cultivation of Science

Kolkata – 700032, INDIA

Email : [ccsd@iacs.res.in](mailto:ccsd@iacs.res.in)

### Key Words

Research Library, Digital Information, Internet Search Engine, File Formats, Preservation.

### Abstract

The steady growth of digital information (Lyman and Hal 2003) as a component of major research collections has significant implications for research libraries. The collection of digital resources as well as access through network by research libraries are increasing day by day. Digital Libraries (DLs) are broadly recognized as a key component in the Global Information Infrastructure (Borgman 2000) - a means of sharing information and knowledge between and among the research, business, government, and educational communities. But how far the research libraries in India adopted and adapting this new technology to realize the “anytime, anywhere” access to information of digital libraries? This article presents an overview of the difficulties facing by the research libraries in India as well as their roles in the present chaotic situation of managing more hybrid and distributed collection of both physical and digital materials.

### Introduction

The main objective of a library is to preserve information to facilitate future access and dissemination of knowledge. Digital information is easy to create, copy and disseminate but very hard to preserve (Lawrence, Kehoe, Rieger, Walters and Kenney 2000). Digital information exist in a wide variety of proprietary formats in the absence of any international standard. Long lives of these resources are at risk due to constant threat from hackers and virus infection. This fluid and unsettled environment presents challenges for a library, especially in regard to the library’s collection management decisions and access strategies. Preservation of digital information through migration needs periodic transfer of digital materials from one hardware/software configuration to another or from one generation of computer technology to a subsequent generation. But it requires huge fund and trained manpower. The difficulties facing by the research libraries in India may be divided into three

categories, namely (i) financial, (ii) technical and (iii) cultural.

The **financial** problem is mainly due to

- (a) rising journal prices
- (b) frequent software/hardware obsolescence
- (c) impart training to existing library personnel to handle digital resources

The **technical** problem is related to the lack of expertise on

- (a) collecting and maintaining digital resources
- (b) developing and maintaining online catalogs and library Web server
- (c) finding scholarly information from “hidden Web”

The following organizational **cultures** may be considered as hindrance

- (a) no full time computer professional in the libraries
- (b) lack of coordination between computer professionals within and outside of the organization
- (c) lack of interest to keep track with the ever changing information seeking behaviour of users
- (d) lack of interest to add values to their services
- (e) lack of interest to interact with users
- (f) lack of interest to utilize physical space of the library

The problems and their probable solutions are discussed in the following sections.

### **Where and How to Find?**

The traditionally understood definitions of library collections and access services have been dramatically changed with the help of Web and Internet and ubiquitous developments in digital technologies. The concept of how readers may discover, gain access to and use information (Case 2002) is also changed due to growing availability of these technologies. The Internet has been adopted as a major part of corporate and government information infrastructure. The development of networked information has brought significant changes in publishing and in information-seeking behaviour of researchers, professionals, general public including children. But widespread network access to digital resources has created a paradox for the library people as well as for the academic and research community. The main reason is that there is no uniform and consistent way to find the materials existing over the Net (DLF 2003). The useful digital resources of serious interest to research and education are difficult to find and impossible to search across, as they exist mainly in isolated pockets. Library people help readers to search and access materials physically kept in the library on the basis of descriptive metadata like catalogs and index cards as created beforehand. But to find a digital

resource library people as well as researchers mainly rely on commercial Internet search engines, which cover academic and scholarly materials poorly. Internet search engines allow for the creation of massive indexes of HTML text based on automatic harvesting of Web sites with unreliable quality of materials provided by commercial sector as well as by unknown and uncredentialed individuals and organizations. Internet search engines cover a small portion of total Web space and favour retrieval of resources based on their own business considerations rather than on the needs of searchers. Typical search for the word 'Iglu' or 'Igloo' gave around 3,00,000 listings through Yahoo and Google searches and almost all are not related with the actual meaning of the word which I wanted to know for my son. Finding relevant information on the Internet take a fair amount of time and once resources are found, determining their reliability is often difficult if the user is not already an expert in the field of the resource under consideration. In fact users expect systems that are navigable in more flexible and adaptable ways than query searching which are badly missing. Generally the search results give just pointers to the content rather than the delivery of the full paper unless one pays a hefty sum to the commercial content providers. They are also susceptible to pop-ups, redirection, virus infection and other dubious practices.

### **Access to Information**

There is a vague idea among many that anything available on Net is free. Access to information is essential for scholarly communication as new knowledge is developed from existing information. But the hard reality is that the access to scholarly information is not easy due to rapidly rising journal subscription rates. Libraries are forced to cancel subscription of some journals necessary for research and education. The growing availability and application of digital technologies have revolutionized scholarly communication but it also accelerated the commodification of information by some content industries. A recent study (Case 2003) revealed that expenditure for serials by research libraries increased 210% between 1986-2001. Research libraries were forced to curtail book purchase to sustain journal collection. So library people frequently face odd questions from researchers when they are asked to pay for downloading some scholarly resources. It is evident that one has to pay substantially to access organized database of digital resources of scholarly information from commercial content providers. The problem has become more complex for online electronic journals. Libraries do not own physical copies of electronic resources and the model has changed from the purchase of physical copies to the licensing of access. Sometimes large commercial publishers offer large bundles of electronic journals with multiple year licences which force libraries to cancel titles from smaller publishers to cover price increase of the bundle. This diminishes competition and increases the market control of the large publishers. While there were 13 major STM publishers in 1998, only seven remained in 2002. This mergers and acquisitions lead to the increase in journal prices in absence of market competition. The research libraries

in India and the libraries in the third world in general are the worst sufferer because of paucity of funds allotted to them for journal collection. Online versions of many journals along with printed copies are being published nowadays. To keep pace with the changing service demands of users the libraries are forced to subscribe both print as well as electronic version of the journal if available. Subsequently they pay much which is very difficult for libraries of a developing country like India. Earlier research libraries used to purchase printed copies of serials and those become physical assets of the library. In digital environment if a library discontinues a particular journal due to some reason then they may not be able to access previously subscribed issues of the journal unless download the same beforehand and maintain its own repositories. But there is copyright issue and it is practically impossible to maintain a digital repository without proper infrastructure and technical knowledge. Shelf lives of print outs of digital resources with a standard laser printer and ordinary office stationery are much less than the normal printed journals. Typical worries about electronic journals (Shiebel 2003) are that those are not indexed and have less prestige. Also majority of users do not want to read articles online. Most e-journal users still print out articles that are judged useful. So a printing format such as PDF is more popular than HTML. Archiving and content control is more problematic for e-journals due to little knowledge about metadata of electronic resources. Electronic journals contained in CDs supplied by publishers are not properly used due to poor maintenance and inadequate infrastructure as well as lack of interest of library staff towards using them.

### **Formats and more formats**

The ever changing nature of formats of digital resources (Lawrence, Kehoe, Rieger, Walters and Kenney 2000) pose a serious problem for the research libraries in managing them. Each of these formats continues to evolve and we often face software enhancements leave unreadable files generated by earlier version. Enhanced software are always resource hungry. It is quite understandable how difficult it is for libraries to upgrade hardware and procure software frequently to get access to digital resources of new file formats. The Semantic Web, once thought as system independent, has become more complex. With so many plug-ins and applets, frames and forms, many Web resources cannot be browsed with earlier version of browsers. Higher versions of browsers are not always compatible with existing hardware. Browser specific Websites make the situation more encumbered. This technology driven development makes the task practically impossible for research libraries to manage digital information due to wide variety of data formats and software/hardware obsolescence.

### **Library as a place**

Many research libraries in India do not have the infrastructure for the readers to access digital resources from the library itself. As a result faculties and researchers are bound to access

digital resources from their homes and offices. Hours of valuable time is being wasted to search useful and authentic digital resources without the help of library professionals. Also library as a physical place for social interaction among researchers are not being used. Recent study on user behaviour with electronic resources (Tenopir, 2003) in USA showed that faculty and students use and like electronic resources and most readily adopt them if the sources are perceived as convenient, relevant, and time saving to their natural workflow. No serious study was done on user behaviour and usage in India. But this is important for capacity planning, systems design, user support and strategic planning (Greenstein and Troll 2000). Many research libraries in India are engaged in digitizing rare manuscripts and artifacts under Traditional Knowledge Digital Library (TKDL) project. But without proper copyright and protection, unscrupulous greedy persons with better technology may pirate these invaluable resources in digital form. Authenticity is a concern for both information providers and information users (Bearman and Trant, 1998). It is difficult for researchers to identify and assess the integrity of a particular source of interest due to easy proliferation of source information on the Internet. Library professionals can help users in certifying the authenticity of the digital resources if accessed from the library itself. There is great confusion among library people regarding choice of materials and tools to digitize its own resources. Proper understanding of metadata in digital environment and their role in accessing digital resources are not known explicitly to many library people.

### **Online Catalog**

Many research libraries do not have catalogs of their holdings accessible online to researchers. But the role of online catalog has changed a lot. Online catalog should be integrated with Web for efficient resource-sharing based on linked catalogs and there should be provision to search across them. Research libraries should collect and catalog free Web resources. But regular maintenance is essential to ensure that links are up to date as many electronic resources available freely on Internet are transient in nature. Online catalogs can be more useful if libraries add some values like providing additional information to the links and indexing or categorizing the links. Using OAI metadata harvesting protocol (Lynch 2001) libraries can deliver information from the “hidden Web” not normally found by Internet search engines. Research libraries should develop or adopt library management software to streamline or transform library operations suitable for research communities.

### **Open Access**

Majority of the research libraries in India are unaware of the initiatives on the transition of journals to open access. Open access is a cost effective way to disseminate and use information. It is an alternative to the traditional subscription-based publishing model where access to the journal is restricted to those who subscribe the journal (either directly or through

their institution). In fact journals should not be expensive as writing, editing, reviewing and typesetting are done free of cost by researchers. Publishers are making profit by controlling access, archiving and marketing. Open access refers to works that are created with no expectation of direct monetary return and freely available online to all interested readers on the public Internet for purposes of education and research. In the open access model there is no problem for the authors to own the copyright in their works. Open access is intended to be free for readers, not free for producers. As peer review is independent of medium and funding model, the open access journals can have the same quality standards as closed access journals. The goal of open access is to generate the least encumbered highest quality scholarly communication. Recent study (Prosser 2003) revealed that open access model can increase the impact factor of a journal. Research libraries can play an important role by publicizing open access journals to faculties. They can also run and facilitate open access servers.

### **New Roles**

Service demands as well as pattern of information use by users are changing as a part of cultural changes in the society. Transformations are also underway in the research libraries specifically in collection management and access strategies. There is ample evidence (ARL 2002) that when libraries make quality content available through the Web, its use increases and it reaches more people within the institution. At the same time many libraries see declining use of some traditional services, such as use of some print materials, circulation transactions, and reference transactions. Although students continue to depend on the library for information resources, they think library services do not meet their expectations for Web-based information services. Users always seek new tools, services, and facilities to access digital resources. Personnel policies and practices as well as budgetary reallocation of the libraries are getting more flexible to incorporate new services to meet the users expectations. Consortia are being formed for the acquisition of electronic content as the whole process is not orderly as print journals. Libraries are expanding the amount and variety of high-quality information resources that are directly available to academic and research users via the Web. Collections also include “born-digital” content that are neither owned nor licensed by the library. Libraries are expanding e-services & tools to enhance user access through online catalogs and library management software. Organizational changes within the libraries to innovate and improve services are underway in many research libraries. Redeployment of staff, imparting training to library staff to perform across traditional roles and to build bridges between the culture of the library and that of the faculty and information technology. Libraries are supporting the development of institution-wide knowledge management systems which might include collections of digital material created by faculty, research staff and students. Less frequently used print resources are being relocated to make room for online access from the library. Portal software are being used that enables customized searching across both

licensed and freely available information resources as well as links to supporting services, such as document delivery. Libraries are being renovated to expand e-access and foster community by providing electronic classrooms, wireless data networks and expanding library hours.

## **Conclusions**

Research libraries should play an important role in the organization and availability of digital information. A uniform policy should be evolved on the line of Open Archives Initiative (OAI) to bring the research libraries in India under one umbrella to facilitate the efficient dissemination of digital resources. Research libraries in India should embrace the initiatives stimulated or endorsed by different Open Access projects (BOAI, SPARC etc) that experiment with alternatives to the current subscription based funding model or the current journal-based publishing model for scholarly communication. Research libraries should encourage faculties and scholars to publish their work where open access is possible. They should monitor trends, assessing their performance by seeking feedback from current users and anticipate the needs of future users. Library and Information Science course curricula should introduce latest techniques in managing digital content to produce library professionals of next generation. Attitude of existing library staff towards the changing service demands of users should be positive. There is no substitute for imparting training to promote the skills and competencies for librarians to carry out the new roles of reaching out to interact with faculty. Libraries can play the important role in certifying the authenticity and provenance of content on the Semantic Web. Research libraries need to work with scholars and scientists to develop scholarly communication systems that serve the educational enterprise rather than exploit it.

## **References**

1. ARL. 2002. **Collections & Access for the 21st-Century Scholar: Changing Roles of Research Libraries**. A Report from the ARL Collections & Access Issues Task Force. *ARL Bimonthly Report* **225** (December 2002).
2. Bearman, David and Trant, Jennifer. 1998. **Authenticity of Digital Resources: Towards a Statement of Requirements in the Research Process** *D-Lib Magazine* (June 1998).
3. Borgman, Christine L. 2000. *From Gutenberg to the Global Information Infrastructure : Access to Information in the Networked World*. Cambridge, MA: The MIT Press. pp 324 including index
4. Case, Mary M and Adler, Prudence S. 2002. **Promoting Open Access: Developing New Strategies for Managing Copyright and Intellectual Property**. *ARL Bimonthly Report* **220**, February 2002.
5. Case, Mary M. 2003. **Framing the Issue : Open Access**. *ARL Bimonthly Report* 226,

February 2003.

6. DLF, 2003. **A New Approach to Finding Research Materials on the Web.**  
URL <http://www.diglib.org/architectures/vision.htm> (Accessed on 10th December, 2003)
7. Greenstein, Dan and Troll, Denise. 2000. **Usage, Usability and User Support.** Report of a discussion group convened at the DLF Fourm on 2 April 2000.
8. Lawrence Gregory W, Kehoe William R, Rieger Oya Y, Walters William H and Kenney Anne R. 2000. **Risk Management of Digital Information : A File Format Investigation.** A Council on Library and Information Resources Publication, (June 2000)
9. Lyman, Peter and Hal, Varian R. 2003. **How Much Information?** Regents of the University of California, October 27, 2003. Available at <<http://www.sims.berkeley.edu/research/projects/how-much-info-2003>>.
10. Lynch, Clifford A. 2001. **Metadata Harvesting and the Open Archives Initiative.** *ARL Bimonthly Report* **217**, August 2001
11. Prosser, David 2003. **On the Transition of Journals to Open Access.** *ARL Bimonthly Report* **227** (April 2003)
12. Shiebel, Stuart M 2003. **The hard Reality of Moving Toward an Open Access Model.** 143rd ARL Membership Meeting, October 16, 2003
13. Tenopir, Carol. 2003. **Use and Users of Electronic Library Resources: An overview and Analysis of Recent Research Studies.** A Council on Library and Information Resources Publication, August 2003