

Digital Library: Challenges and Issues Pertaining to Preservation

Hemanta Kumar Gohain

Abstract

One of the most frequently quoted word in the field of library and information science in recent time is digital library. Advances in computer storage and telecommunication methods, online access to databases, electronic journals, electronic document delivery, Radio Frequency Identification systems, Compact discs, Internet etc are revolutionary developments in the last two to three decades that have brought a great change in the collection, communication, storage and management of information. The present scenario reflects the emerging trends among libraries to get themselves into digital form with a vision to build up a sound digital resource collection by transforming from traditional to digital. The role of Information and Communication Technology in library and information sector is enormous and its impact has compelled the traditional libraries to revisit their modus operandi in order to deliver speedy access of information to their client. Skill that need for establishing digital library becomes basic necessity for capacity building in respect of not only collection development but also services to be rendered. The main purpose of digitization is to improve the efficiency of Library and to render optimum user service, which enables a user from single workstation to perform wider variety of activities. Digital preservation is the most effective methods through which we can permanently preserve any information specially those are of national importance. Before taking the final decision on digitization, it is better to determine the main purpose of the digitization whether digitization is primary for preservation or for frequent use. Digital Library has its own merits and demerits. Digital Libraries are engaged to collect, store and disseminate digital information. . In digitization process focus should be given feasibility as we cannot digitize all the printed materials even we do it may not be cost effective. Most important is that what to digitize and so the priority.

Keywords: Digital Library, Internet, ICT, Hybrid Library, Web Server, Digital Preservation

1. Introduction

In 1978, Federick Wilfrid Lancaster put forward the idea of paperless society. A Paperless society is one in which paper communication (written documents, mail, letters, etc.) is replaced by electronic communication and storage. Forecasting the role of Librarian in such environment Lancaster opined that “Librarians will, in time, become information specialists in a deinstitutionalized setting” (Lancaster, 1980). The concept of Digital Library was conceived in the later part of 1980. It took only few years to establish and now it is a global phenomenon. . Ad-

vances in computer storage and telecommunication methods, online access to databases, electronic journals, electronic document delivery, Radio Frequency Identification systems, Compact discs, Internet etc are revolutionary developments in the last two to three decades that have brought a great change in the collection, communication, storage and management of information. As a result, the organization, structure and functions of the libraries have changed and as such the type of libraries divided into three categories such as Conventional Libraries, Digital Libraries and Hybrid Libraries. The Conventional or traditional libraries deal with printed collection. Digital Libraries are engaged to collect, store and disseminate digital information and the Hybrid Library



co-exists with both print and digital collection and its services. The present scenario reflects the emerging trends among libraries to get themselves into digital form with a vision to build up a sound digital resource collection by transforming from traditional to digital. The role of Information and Communication Technology in library and information sector is enormous and its impact has compelled the traditional libraries to revisit their modus operandi in order to deliver speedy access of information to their client. While doing so it is necessary to ensure whatever digital transformation is in agenda, predominately it must comply with the public demand and pressure from the user. To acquire, convert, process and maintain such resources need specialised knowledge, competence, commitment and management skill. Therefore, skill, which is needed for establishing digital library, collection development / management and to render services, becomes the basic necessity for capacity building.

The National Leadership Grants Program of Institute of Museum and Library Service (U.S.A.) stated that 'the digital revolution has affected nearly every aspect of library and museum services, from automation of internal recordkeeping system to the digitization of physical collection and from acquisition of new "born digital work" of arts or library publication to the use of technology to present collection and engaged audience. The digital technology enables the full range of holdings in our museums, libraries and archives- audio, video, print, photograph, artworks, artifacts and other resources- to be catalogued, organized, and combined in new way and made accessible to audience in new way. Digital Technology connects more people to the resources and services only museum and library can provide.' (IMLS, 2005).

2. Digital Library

One of the most frequently quoted word in the field of library and information science in recent time is digital library. The digitization means the conversion of any fixed or analogue media such as books, journals, articles, photos and paintings into electronic form through scanning, sampling etc. where they can be stored and manipulated by a computer (Kaur,2009). The term Digital Library in broad sense is automated system that allows users to obtain a coherent means of access to an organized electronically stored repository of information and data. One of the most popular definition by Marchinini (1999) which defined that the digital libraries are the logical extensions and augmentation of physical libraries, and that digital libraries are distinguished by focus on integration of services through a holistic treatment of interface location, time and language and system. There is enamours literature on digital library which gives new dimension on different issues pertaining to digital library. Usually such definitions convey different concepts. However, there is no single agreed upon definition of what constitutes a 'digital library' (Spink and cool, 1999). Among them the definition given by Digital Library Federation has drawn substantial attention by LIS professional. Digital Library Federation defined "digital libraries are organizations that provide the resources, including specialised staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collection of digital works so that they are readily and economically available for use by a defined community or set of communities" (DLF)(1998).

3. Why Digital Library

Whenever we think about changes in library services, focus should be always on users. The ICT

has made tremendous changes in information sector and library is also not an exception to it. It is now a universal fact that the users' information seeking behaviours has been gradually shifted from printed to digital resources because of convenience and easy to access. Therefore use of printed material has declined sharply over the years. Today's digital world is connected with creating, sharing and using information in digital form because the right information to the right person at right time at fastest possible way is the basic expectation of modern information user. Moreover services are also expected beyond the physical entity of the library and Information centre. To meet the new expectations of information users, we have no option but to adopt appropriate information and communication technology to shift from print to digital. Advancement in digital technology and increasing application of these technologies in libraries has brought in a revolutionary paradigm shift in all types of libraries. This shift can be identified as:

From	To
Custodian of Book	Service oriented information provider
One medium	Multiple media
Own Collection	Library without wall
In good time	Just in time
In- sourcing	Out- sourcing
Local reached	Global reach
User goes to library	Library comes to user

Digital library provides such information service in which all the resources are available in digital format and the functions of the library such as acquisition, preservation, retrieval, access and display are carried out through digital technologies. Under digital library system in addition to full text access user can also access non text information like photographs, drawing, illustration, artwork, numeric data, cosmological data, GIS, digital sound, moving vir-

tual images and also 3D representation. In digital library system information in any format, whether textual, images or any other form can be disseminated through powerful network system and user can access the same anytime and anywhere in the world. The main benefits of digital library are:

- ❖ Provide fast access to very large information with multiple access point.
- ❖ Provide solution to space
- ❖ Support multimedia content
- ❖ Preservation is added advantage
- ❖ Provide user friendly interface to access the information
- ❖ Time and energy saving in finding information
- ❖ Provide remote access
- ❖ Support advance search and retrieval
- ❖ Comparatively economical in long run
- ❖ Easy identification of relevant document from millions of records
- ❖ Could be interactive
- ❖ Availability of pre-print copy before publication thus reducing time lag.
- ❖ Automatic Alerting Services (User oriented service, interest profile, Selective Dissemination of Information etc)

Apart from the above benefits popularity of digital library is increasing because of its unique characteristics. The USA Association of Research Libraries identified five unique characteristics of digital library:

- ❖ Digital Library is not a single entity
- ❖ Digital Library requires technology to link the resources.
- ❖ Linkage between digital library and Information services are transparent
- ❖ Universal Access to digital library must be a goal
- ❖ Digital Library collection are restricted to document surrogates but includes digital artefacts that have no printed equivalent

4. Planning of Digital Library

The main purpose of digitization is to improve the efficiency of Library and to render optimum user service, which enables a user from single workstation to perform wider variety of activities. Prior to planning it is necessary to decided whether the establishment of digital library is based on policy matter of the authority or its creation is consequence of public demand and pressure from the user community. First one is a routine affair and second one needs to be addressed from the point view of information need of the user community.

Infrastructure is the basic requirement to lunch any digitization project. The planning would depend upon the nature and scope of the project. However, principle of planning mostly remains same which should be done for next five year if not more.

4.1 Infrastructure Requirement

Process of digitization mainly includes scanning (image creation), merging, splitting, input, indexing and storage. Building of suitable infrastructure needs skill, and that can be achieved through experience, training and appropriate consultancy. Sound policy is very much essential for infrastructure building.

Many experts on digitization opined that the consultation in India is neither easily available nor sought seriously. There are many instances where money has been spent generously but not interested in spending on consultancy and even if spend the amount is inadequate. In such a case the planner has to visit well established digital labs and prepare a detailed list of infrastructure available in consultation with staff engaged in digitization process. It is always useful to learn from other's experience and expertise. The preliminary list so prepared should be discussed thoroughly with other expert group and also make detailed review to check whether it fulfils the requirement of parent organization before taking the final decision.

4.1.1 Servers

Digital Library would need three servers:

- ❖ Web Server
- ❖ Application Server
- ❖ Database Server
- ❖ With SAN Storage

Blade servers are now preferred having high storage and data processing capacity depending upon the quantum of data likely to be processed and input. The RAM of server is needed in GB whereas SAN storage should preferably be in TB. The capacity of hard disk is recommended for more than 500 GB as PDF file or other formats of files which go into digital library are more space consuming.

4.1.2 Works Stations

The number of workstations would depend upon the size of the network going to be planned and the number of the staff likely to work. But in general each library that undertakes digital library project

would require at least six workstations to take up scanning, data input and data editing simultaneously.

4.1.3 Digital Library Software

There is variety of software used for digital library. However there are some useful software available in open source which can be easily downloaded and installed and have succeeded in drawing attention of the library professionals. Some of them in use are:

- ❖ DSpace: This has been developed by MIT and is being distributed by HP
- ❖ Greenstone: This Software was developed out of New Zealand Digital Library Project at the University of Waikato being distributed in cooperation with UNESCO and Human Info NGO
- ❖ E- Print: This software is most frequently used and widely distributed all over the world. This was developed by the University of Southampton.
- ❖ Archmede and CD Sware (CERN Document Server Software) are other commonly used software developed by Lavel University Library Canada and The European Corration for Nuclear Research respectively.

All the above software have their own merits and demerits hence proper evaluation is need before final choice is made. However, in India DSpace has perhaps maximum installation and is being used successfully.

4.1.4 Network Requirement

High speed bandwidth of internet connectivity is one of the most essential perquisites for effective

digital library system. This is required for data points and their connectivity with the main server of the institution controlling the internet connectivity. These are minimum requirement of establishing digital library.

4.1.5 Scanners

Digitization starts with image making, hence selection of scanner is very important for digital library. The decision has to be taken according to the requirement of present and future project to be taken up. The technical specification need to be first worked out according to the requirement and then match with available models in the market. Some of the basic features need to be evaluated are: optical resolution, speed, image type, paper thickness acceptability, software features, file formats, scanning mode etc

5. Artitechure

- ❖ Data capturing
- ❖ Image preparation of document
- ❖ Scanning of pages
- ❖ Editing of scanned pages including merging and splitting of files
- ❖ Use of OCR Technology to edit text and convert it into ASCII text
- ❖ Creation of metadata (including indexing)
- ❖ Linking text/ video/audio files to metadata
- ❖ Follow of work flow (accepting, approving, editing and archiving)
- ❖ Searching and retrieval
- ❖ User interfaces, users profiling services
- ❖ User authentication and authorizing services

6. Formal Planning

Before launching a digitization project, digitization committee should be constituted comprising of members from various fields like from library, administration, members from technical group, users. Moreover external member with experience in digitization will be quite useful. The following important issues need to be addressed properly while making the final shape:

- ❖ Costing of Project
- ❖ Assessment of use
- ❖ Selection of materials
- ❖ Copy write management
- ❖ Preparation of management
- ❖ Conversion
- ❖ Adding metadata
- ❖ Retrieval aspects
- ❖ Making provision for archiving and preservation, etc.

7. Digital Preservation

In modern era it is fundamental duty of the nation to preserve all the important information that is of historical importance. Digital preservation is the most effective method through which we can permanently preserve any information, especially those are of national importance. So far as digital archiving is concerned it is relatively a new discipline though, microfilming and microfiche technology are quite old and still are in use with the advantage and disadvantages of its technology. Digital preservation combines policies, strategies and actions that ensure access to information in digital formats over time (ALCTS, 2007).

7.1 Objectives of Digital Preservation

The objective digital preservation can be summarised as under:

- ❖ “Each item in the archive is quality assessed and functional to fullest extent by current technical capabilities.
- ❖ A gathering schedule can be individually tailored for each selected title taking into account its publication schedule or the frequency with which the website changes, thus enabling the content gathered to be as complete as possible.
- ❖ Each item in the archive can be fully catalogued and therefore, can become part of national bibliography.
- ❖ Each item in the archive can be made accessible via web to reader immediately because permission to do so can be negotiated with the publisher.
- ❖ The “significant properties” of individual resources and classes of resources can be analysed and determined.
- ❖ Sites that are inaccessible to harvest robots can be identified and archived using other methods, by arrangement of other methods with the author.”

(Phillip, 2005)

.Before taking the final decision on digitization, it is better to determine the main purpose I.E. whether digitization is primary for preservation or for frequent use. If primary purpose is preservation we may consider other technology like microfilm as it has life expectancy of 500 years whereas digital resources are more susceptible to be lost accidentally

or due to change of hardware and software. In the changing condition of hardware and software, it needs proper maintenance and timely migration from one format to another (Chandel, A and Mannan S. ,M 2003).

Normally there are two types of materials selected for digital preservation one is in the public domain and another is someone other than the institution owns the copyright. Communication with copyright owner and licence holder before preservation work is very much essential to avoid violation of copyright law. Copyright act provides exclusive right to copy an item of work to the author/copyright holder. In addition he also enjoys following rights:

- ❖ Prepare directive works based upon the copyrighted work.
- ❖ Distribute copies of copyright work to the public.
- ❖ Perform some copyrighted works publicity.
- ❖ Display some copyrighted works publicity in the case of sound recordings, to perform the copyrighted work publicity by means of a digital audio transmission in United States.
- ❖ Control access to a work protected by the use of a technological measure.

Under such circumstances, the libraries and archives can restrict their preservation activities to carry out the task with fear of infringement. However following requirements apply:

- ❖ Copy of the original must be obtained.
- ❖ The copying must be solely for preservation or security.

- ❖ The original must be “damaged, deteriorating, lost, or stolen“ or the existing format in which the work is stored is obsolete.

A reasonable investigation reveals that an unused copy cannot be obtained at a fair price (Chadha,R. K., Chandel, A. 2003).

7.2 Material to be Digitised

1. Institutional Repository (IR)
2. Rear collection which are usually available in single copy and are in danger being lost in near future; such as old manuscript, old photograph, old music recordings, newspapers, cultural artefacts etc
3. Creation of special database; like children collection, collection for disabled, endangered species of plant and animals.
4. Information on indigenous knowledge especially on sustainable development.
5. Local history collection /country history
6. Digitization of important monograph not under the purview of copy right but often referred.
7. Archive of ethno-musicological videos that capture the multitude of cultural practices-including costumes, ritual practise and dance – that are integral to fully understanding musical expression. (Chandel, A and Mannan S. ,M 2003).

The use of the proposed digital resources should be open for both local as well as global level. Preference should be give to users for whom the digital resource is created and the local community. Hence the priority of library will defer from one another as

they serve variety of users. Digitization should be cost effective and output of the digitization must have the scope of intensive use.

7.3 Repackaging of some Selected Digital Material

There is certain information which is already available in digital form but needs to be repackaged in more presentable way for easy access. Identification, collection and organization of such materials is an essential activity of an effective digitisation project. This would require capturing of such resources, creation of their metadata and uploading of relevant file. This would skip the scanning stage which is required to convert printed material to digital.

8. Conclusion

Digital Library has its own merits and demerits. Digital Libraries are engaged to collect, store and disseminate digital information. Due to the advent of ICT the Library and Information Centre have to rethink about the style of functioning as users' demand for faster dissemination of information is increasing day by day. Services provided by digital library can give solution to this dimension. While doing so focus should be on feasibility as we cannot digitize all the printed materials even if we do it may not be cost effective. Most important is what to digitize and so the priority. The best solution could be to conduct feasibility study and find users' preference and possibility of present and future usability. Advisory body on digitization comprising experts as well as other stake holder of the system and preparation of policy document explaining principle and guidelines of digitization may prove helpful.

References

1. ALCTS. Preservation and reformatting selection committee (PARS). Digital Preservation group blog [Available at : [http:// blog.ala.org/digipres.php](http://blog.ala.org/digipres.php)
2. Chadha, R. K., Chandel, A (2011). Knowledge Management in Digital Era. Edited by A C Chandel, New Delhi: Westville Publishing House, 214,
3. Chandel, A and Mannan S., M (2011). Knowledge Management in Digital Era. Edited by A C Chandel, New Delhi: Westville Publishing House, 179-190,
4. Chandel, A and Mannan S., M (2011). Knowledge Management in Digital Era. Edited by A C Chandel, New Delhi: Westville Publishing House, 182-183,
5. Institute of Museum and Library Services (IMLS) (2005). National initiatives: Digital corner. US federal government. **In** Seadle, Michael and Greifeneder (2007). Defining digital Library(editorial) Library Hi Tech, 25(2), 169-70
6. Kaur, S (2009). Digitizing Library Resources. Digitization of Library and Information Centre in North East Indi. Edited By N. Lahkar and S. K. Singh, Guwahati: Department of Library Science, 241-245, 2009
7. Lancaster, F. W. (1980). The future of the librarian lies outside of the library. Catholic Library World, 51, 388-391
8. Marchionini, G (1999) Overview of digital libraries. School of Library and information Science. University of North Caroline. **In** Digital Libraries in Public libraies/christinger Tomar. Encyclopedia of library and Information Science. 2 nd edited by Miriam A Drake. New York: Marcel Dekker, 884, 2003

9. Phillips, Margreat, E (2005). What should be preserve? The question for heritage libraries in digital world. Library Trends, 54(1), 61-62
10. Spink, A and Cool, C (1999). Education for digital Libraries. D Lib Magazine, 5. **In** Digital libraries in public libraries/ Christinger Tomar. Encyclopedia of Library and Information science. 2nd. / edited by Miriam A Drake. New York: Marcel Dekker.884,2003

About Author

Mr. Hemanta Kumar Gohain, Technical Assistant,
Assam Agricultural University, Jorhat (Assam)