



SIGNIFICANCE OF DIGITAL LIBRARIES: AN OVERVIEW

YUGAL JOSHI

Information Scientist,
Kumaun University,
Nainital
ycjoshi@yahoo.com

Abstract

We have seen several recent development initiatives of digital library in India. In order to enlighten, assess and understand the growth, development and current status of digital library initiatives in India as reflected through scholarly review journals, 20 published studies on digital libraries in India have been examined. The paper discloses that most articles focus on developing digital libraries and digital collections except for a few studies on copyright issues and management of digital libraries. No studies have touched upon issues such as digital rights management, security and digital library policies.

Keyword: Digital library, Digital rights management, Digital library policies and Security, Dublin Core, MARC, Metadata Object Description Schema, Text Encoding Initiative (TEI), Encoded Archival Description (EAD) etc..

Introduction

The Information and Communication Technology (ICT) has changed radically the way people think and work. It has transformed the functioning of individuals and institutions. The telecommunications, mobile technology, email, social media have brought in a new wave of change in all walks of life. World Wide Web has become the essential infrastructure of day-to-day life in the 21st century. The digital information is the most important source of information created in this century. Digitization or creation of information in digital form has enabled all the components of the information chain to be most dynamic. The authors, publishers, libraries and other information service providers are playing a new role. They are further getting integrated in to a great extent in this chain. The new roles have resulted in the components of the information chain being converted in to digital artifacts. The knowledge contained in these digital artifacts must be interpreted in the social contexts within which they are published and used. Practically, a new public realm must integrate speech, print and digital expression. The physical features of the digital artifacts within which knowledge is contained, whether books or computers, will shape the dynamics of reading and writing, intellectual property markets and lastly the shared modes of learning and knowledge. Thus, a book and a digital resource are different kinds of knowledge artifacts and the differences will have impact on the way ideas are conceived and implemented, but they are not contradictory. The digital collection in various formats, available in many libraries call for developing the crucial digital competencies, resulting in new challenges being faced by library users. One of the challenges is how well the relevant information that is available can be accessed and secondly, what is the best way to use the digital resources to serve the academic purpose. The teaching community and the librarians need to integrate digital resources in the classrooms for better understanding by the students. This concern calls for the basic understanding of the various perspectives of digital libraries in the context of the academic arena.

History of Development of Digital Libraries

The early history of libraries is poorly documented, but several key thinkers are connected to the emergence of this concept. Predecessors include Paul Otlet and Henri La Fontaine's Mundaneum, an attempt begun in 1895 to gather and systematically catalogue the world's knowledge, the hope of bringing about world peace. The establishment of the digital

library was total dependent on the progress in the age of the internet. It not only provided the means to compile the digital library but the access to the books by millions of individuals on the World Wide Web.

Vannevar Bush and J.C.R. Licklider are two contributors that advanced this idea into then current technology. Bush had supported research that led to the bomb that was dropped on Hiroshima. After seeing the disaster, he wanted to create a machine that would show how technology can lead to understanding instead of destruction. This machine would include a desk with two screens, switches and buttons, and a keyboard. He named this the "Memex." This way individual would be able to access stored books and files at a rapid speed. In 1956, Ford Foundation funded Licklider to analyze how libraries could be improved with technology. Almost a decade later, his book entitled "*Libraries of the Future*" included his vision. He wanted to create a system that would use computers and networks so human knowledge would be accessible for human needs and feedback would be automatic for machine purposes. This system contained three components, the corpus of knowledge, the question, and the answer. Licklider called it a precognitive system.

Early projects centered on the creation of an electronic card catalogue known as Online Public Access Catalog (OPAC). By the 1980s, the success of these endeavors resulted in OPAC replacing the traditional card catalog in many academic, public and special libraries. This permitted libraries to undertake additional rewarding co-operative efforts to support resource sharing and expand access to library materials beyond an individual library.

An early example of a digital library is the Education Resources Information Center (ERIC), a database of education citations and abstracts, which was created in 1964 and made available online through DIALOG in 1969.

Digital Library: Meaning and Definition

Digital Library represents different meaning to different people and organizations. A Digital Library may mean a collection of digital documents, database, video games and learning materials accessible via computer network, to children. The collection that may be available over Internet in the form of GIS and CAD data, satellite imagery, video gallery may mean a digital library for a space scientist. For a businessman, digital library is a collection of information over a portal in the form of important business deals, stocks and shares, budget information, etc. In simple terms. Digital Library is a collection of information which is digitized, organized for a group of people or community, gives users power they had never with traditional libraries. There are terms used in various contexts for use of the digital library, for example, referring to remotely related activities like multimedia database, information mining, information warehouse, information retrieval, on-line information repositories, electronic library, virtual library, image applications, digital preservation, digital archive, publisher databases, e-Journals, eBooks, etc.

Digital libraries can be defined to represent two lines of thoughts. One concentrates on access and retrieval of digital content, which results from the work of computer scientists, engineers, researchers, etc. The second line focuses on the collection, organization and service aspects of digital libraries which can be considered the contribution of library and information professionals practicing at the ground level. Stanford Digital Library research team defined digital libraries "as a coordinated collection of services, which are based on collections of materials, some of which may not be directly under the control of the organization providing a service in which they play a role." (Reich and Winograd, 1995)

Changing Phases of Digital Libraries

In the course of five decades of development of digital libraries, it is observed that there are three distinctive phases in the entire progress of the digital libraries. The first and the foremost is the emergence of online indexing and abstracting periodicals such as Biological Abstracts, Mathematical Reviews, Chemical Abstracts and Engineering Index etc.

The second phase is the dawn of CD-ROMs and the Graphic User Interface, user friendly search software facilitating designing the dynamic databases whereas the third phase is noticed as birth of Internet and WWW which have outscored all time record and changed the entire scenario and gave new direction and dimension to handling information. This has been further elaborated below.

- **Online Technologies**

The first online database suitable for searching was developed in the early 1960s in spite of the advent of the computers in the 1950s. The MEDLARS was the first on-demand computer based information retrieval service developed basically for the medical profession. MEDLINE was the first major online dialup database search service in the year 1971. DIALOG offered the first public online commercial database. Around the same time, INIS AtomIndex was introduced to cover nuclear science literature. Libraries charged the users for searching of these databases. Electronic resources started having an impact on selection practices with the introduction of the CD-ROM in the mid 1980s.

- **CD-ROM Technologies**

Versions of larger online databases were offered to libraries in the form of CD-ROM products initially and were supplied on a subscription basis with ownership of the data remaining with the publisher. The price of the product included licensing of the content and also acquiring the computer and CD-ROM player. As standards were not yet established, products were guaranteed to work only with specified CD-ROM players. The purchase of the equipment as part of the cost of the information product was not appropriate use of the library's materials budget. As personal computers became popular in the libraries, the CD-ROM products became affordable. The major drawback was that only one person could use these CD-ROM databases at a time. The other option to provide simultaneous access to many users was by mounting the needed database on the local computer system was an expensive alternative which many libraries could not afford. With the advent of better hardware and software solutions, libraries started providing simultaneous access to the same CD-ROM database, even to the sites outside of the library.

- **Web Technology**

These technological advances changed the scenario and the librarians quickly managed to handle this new situation and they continued to make thoughtful decisions for getting access to or owning these costly products. The selection decisions were taken by subject specialists, reference librarians, instruction librarians and technical staff. The onset of the latest technology - that of the World Wide Web put the libraries in a spin as they had just managed to get a grip over the selection of electronic resources on CD-ROM. Now most of the bibliographic and full text of databases is accessible through the web. Web has become a platform for publishing and accessing the e-resources.

The World Wide Web has taken over all aspects of computing. This is the preferred media by the information producers to distribute their products and services, at present. A revolution is being witnessed in the information landscape with the advancement in electronic information, development of networked access and delivery of new library services.

Major Issues Involved in Digital Libraries

Designing and development of digital libraries is very complex process. It has multi dimension involves numbers of things to look at. Different set of technologies, tools, standards, software, file formats, access and retrieval, and several other things need to take in to account. Therefore to build a digital library, one has to look for appropriate technological solutions and handle the following major aspects and related issues:

- Digital Library Standards
- Information Resources Organisation
- Metadata Standards
- Digital Archiving and Preservation
- Digital Library Services

These have been further elaborated in the following section here;

- **Digital Library Standards**

The very purpose of developing a digital library is to provide wider and seamless access, preserve the content for future use, interact with another similar digital library etc. To ensure this, there is a greater need for adopting various standards and best practices to build interoperable digitallibraries. Some of most important standards used in the digital libraries are listed below:

- **User Interface** - Common web browser compatible to all platforms
- **Data Handling and Interchange** - Graphic Formats - JPEG, TIFF, GIF, PNG, Group 4 Fax, CGM - Structured Documents - HTML, XML, PDF - Moving Pictures/3-D - MPEG, AVI, GIF89A, QuickTime, Real Video, ViviActive, VRML etc.
- **Metadata** - Resource Description - Dublin Core, METS, MODS, - MARC, TEI Headers, Other Open Source and Domain Specific Standards, - PREMIS {Preservation Metadata: Implementation Strategies} - Resource Identification - URN, PURL, DOI, SICI
- **Search and Retrieval** - Federation and Harvesting: FTP-enabled, OAI-PMH for intermittently transfer data from one system to another - Federated search: Z39.50 protocol, SRW Protocol
- **Security, Authentication and Payment Services** - Emerging e-Commerce Standards.

Information Resources Organization

When we have too many items as part the digital collection, to bring together related items a classification systems are used. Traditional libraries have been arranging and organising printed materials such as books, journals on related subjects together and by following modern and widely used library classification schemes. Classification uses notation symbols,

while Cataloguing creates document surrogates which facilitate browsing and search facilities using author, title, series and other elements. These classification schemes have also been used for organising web based /digital information resources according to the disciplines, specific subjects and topics within a discipline. Having been familiar with organisation of print world in our libraries, users find it easy to get information from organised resource structures. Following are some of the web based systems/digital libraries where library classification systems have been followed;

- BUBL LINK, <http://www.bubl.ac.uk/link>. DDC & LCSH
- CyberDewey. <http://www.anthus.com/cyberdewey/cyberdewey.html>.
- SCORPION. <http://orc.rsch.oclc.org:6109/>.
- Cyber Stakes. <http://www.public.iastate.edu/~cyberstack/>.(LC).
- INFOMINE. <http://infomine.ucr.edu>. (LCSH)
- BIOME. <http://www.biome.ac.uk>

However, classification schemes are unable to keep up the pace as digital libraries deal with many new and nascent subjects and formats of documents. Moreover, it is a costly affair to classify and catalogue according to traditional schemes by experts. Various new metadata schemes have been developed to organise disparate digital resources.

Metadata Standards Metadata, in simple terms is data about data. But specifically. Computer Science, Library and Information science communities have adopted the term "metadata" for describing electronic data. There are three categories of metadata which can apply to objects in a digital library. They are

- **Descriptive metadata** describes a resource for discovery and identification; it includes elements like title, abstract, author, and keywords.
- **Structural metadata** indicates how compound objects are put together, for example, how pages are arranged to form chapters.
- **Administrative metadata** provides information to help manage a resource, such as when and how it was created, file type and other technical information, and who can access it. There are several subsets of administrative data.

Two more metadata types listed as separate metadata types are:

- Rights Management Metadata, which deals with Intellectual Property Rights,
- Preservation Metadata, which contains information needed to archive and preserve a resource. There are large numbers of metadata standards available for describing various digital objects.

In order to consistently describe these different types of metadata in the digital library, following a well-accepted standard becomes necessary. Considering different types of digital objects, requirements of different user groups and purpose several new metadata standards have been designed. Depending on one's requirements, one can decide to adopt the metadata standards. Some of standards which are known and widely use are described below:

- Dublin Core is a list of 15 basic fields designed initially to describe web-based resources sufficiently to allow their discovery by search engines.
- MARC is the established standard for the creation of machine readable cataloguing records, and underlies virtually all online library catalogues.



- Metadata Object Description Schema (MODS) An XML based descriptive metadata standard as a derivative of MARC
- Text Encoding Initiative (TEI) is the de facto standard for the encoding of most types of electronic texts, and as such is used by almost all of the world's e-text centers.
- Encoded Archival Description (EAD) is an XML DTD used throughout the archival community for the encoding of finding aids (collection-level descriptions).
- PREservation Metadata Implementation Strategies (PREMIS) intended to ensure the long-term usability of a digital resource.

Digital Archiving and Preservation

Archiving and preservation of digital information for future use is one main criteria for this the digital libraries are created. Since this aspect involves technologies associated with preservation, file formats, standards and platforms on which digital collection is hosted, it calls for well thought out long term plan and strategy. Unlike the print world, it is difficult to ensure that the digital information is not lost. The digital information may be lost for many reasons. Some of reasons could be because of changes in an organisation, due to content reorganization, cession of sponsorship or support, technology obsolescence, content format obsolescence, hacking and sabotage of data and data files, disaster, whether nature or manmade and other unknown factors. (Stewart, 2000) Now that the many digital libraries have been created and been around for more than a decade, it is important to look at the best practices and tools and technologies used, before one ventures in to creating digital library.

Digital Library Services

Digital libraries are created to break the existing barriers posed by traditional libraries and offer many new services following the spirit and letters of Five Laws of Library Science propounded by Dr. S. R. Ranganathan. The services like reference and information services have been integral to traditional services. Reference services should form an important part of digital libraries, but digital library research and development has not concentrated on them much. A number of free and fee based reference and information services are now available through web like. These are

- Ask Jeeves
- Ask A Librarian
- CDRS- - Automatic Reference Librarians for the World Wide Web.
- Virtual Reference Desk
- QuestionPoint from OCLC
- LibrariH3LP
- AskCollarado

In addition, several other services one can consider adding to digital libraries to promote the access to collection available.

Drawbacks of digital libraries



Digital libraries, or at least their digital collections, unfortunately also have brought their own problems and challenges in areas such as:

- User authentication for access to collections
- Copyright
- Digital preservation (see above)
- Equity of access (see digital divide)
- Interface design
- Interoperability between systems and software
- Information organization
- Inefficient or non-existent taxonomy practices (especially with historical material)
- Training and development
- Quality of metadata
- Exorbitant cost of building/maintaining the terabytes of storage, servers, and redundancies necessary for a functional digital collection

There are many large scale digitization projects that perpetuate these problems.

Future development

Large scale digitization projects are underway at Google, the Million Book Project, and Internet Archive. With continued improvements in book handling and presentation technologies such as optical character recognition and development of alternative depositories and business models, digital libraries are rapidly growing in popularity. Just as libraries have ventured into audio and video collections, so have digital libraries such as the Internet Archive. Google Books project recently received a court victory on proceeding with their book-scanning project that was halted by the Authors' guild. This helped open the road for libraries to work with Google to better reach patrons who are accustomed to computerized information.

References

1. Bush, Vannevar (July 1945). "As We May Think" (PDF). *The Atlantic Monthly*: 101–108. Archived from the original (PDF) on 21 April 2018. Retrieved 30 April 2018.
2. Bourne, Charles P.; Hahn, Trudi Bellardo (2003). *A History of Online Information Services, 1963–1976*. MIT Press. pp. 169–170. ISBN 9780262261753. Retrieved 30 April 2018. 1696 Milestone - DIALOG, with the ERIC database, provided the first instance of extensive availability of abstracts online for search output.
3. Chowdhury, G. G. (2003). *Introduction to Digital Libraries*. London: Facet Publication.
4. Christine L., Borgman (1999). What are digital libraries? Competing visions. *Information Processing & Management*, 35(3), pp. 227-243.
5. Engelbart, D. (1962). *Augmenting human intellect: a conceptual framework* (Summary Report AFOSR-3233). Menlo Park, CA: Stanford Research Institute.
6. Fox, E. A. & Urs, S. R. (2002). Digital libraries. *Annual Review of Information Science and Technology*, 36(1), pp. 502-589.



7. "Fair Use Triumphs as US Supreme Court Rejects Challenge to Google Book-Scanning Project". 2016-04-18.
8. Fox, E. A. (1999). *Digital Libraries Initiative (DLI) Projects 1994-1999*. *Bulletin of the American Society for Information Science and Technology*, 26(1), pp. 7-11.
9. Granger, S. (2000). *Emulation as a Digital Preservation Strategy*. *D-Lib Magazine*, 6(10).
10. Ian H., W. [et. al.,] (2010). *How to build a Digital Library*. New York: Morgan Kaufmann.
11. Licklider, J. C. R. (1965). *Libraries of the Future*. Cambridge: MIT Press.
12. Lynch, Clifford (2005). "Where Do We Go From Here? The Next Decade for Digital Libraries". *D-Lib Magazine*. 11 (7/8). ISSN 1082-9873. Archived from the original on 16 Jan 2018. Retrieved 30 April 2018. This is a field with an incredibly rich, and, as yet, poorly chronicled pre-history and early history. There is a stream of work and ideas that reaches back to at least the turn of the 20th century, and includes such thinkers as H.G. Wells and Paul Otlet; later contributors to the pre-history of visions of new, technologically-enabled means of knowledge organization, access and distribution also include Vannevar Bush and J.C.R. Licklider.
13. Reich, Vicky & Winograd, Terry (1995) Working assumptions about the digital library. Technical Report. Stanford InfoLab. <http://dbpubs.stanford.edu:8090/pub/1995-52>
14. Robinson, John D. (2015). "The dogs bark and the circus moves on" (PDF). *The Bottom Line*. 28 (1/2): 11–14. doi:10.1108/BL-01-2015-0002.
15. Seaman, David, (n.d.). Aggregation, integration and openness: current trends in digital libraries. In *International Symposium on Digital Libraries and Knowledge Communities in Networked Information Society (DLKC'04)*.
16. Sujin, B. (2003). *Basic issues in digital libraries: The 21st century knowledge centers*. Hyderabad: Professional Books Publishers.
17. Stocker, Gerfried (1 January 2014). "Beyond Archives (or the Internet 100 years before the Internet)". In Magalhães, Ana Gonçalves; Beiguelman, Giselle (eds.). *Possible Futures: art, museums and digital archives*. ISBN 9788575963548. Retrieved 30 April 2018. Actually it was 1895 when Paul Otlet together with Henry La Fontaine, who was later awarded the Nobel Peace Prize, started a project - Mundaneum - that was initiated and driven by their idea that, if they would be able to collect all human knowledge and make it accessible to everybody worldwide, then this would bring about peace on Earth.
18. Slideplayer.com Fetched: 2019-10-01T21:33:00
19. Url: <https://slideplayer.com/slide/716833/>
20. Weeks, A. C. (2003). *The International Children's Digital Library*. Presented at the 69th IFLA world Library and Information Congress 1-9, Berlin.
21. Wells, H. G. (1938). World Brain. In Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/World_Brain
22. www.ijsr.net Fetched: 2019-10-01T21:33:00
23. URL: <https://www.ijsr.net/archive/v7i12/ART20193737.pdf>