Chapter 4

Digital Libraries: An Overview
CHAPTER 4
DIGITAL LIBRARIES: AN OVERVIEW

“The scholarship that is being produced today is at serious risk of being lost forever to future generations.”
- Anon

4.1 Introduction

The Information and Communication Technology has changed radically the way people think and work. It has transformed the functioning of individuals and institutions. The telecommunications, mobile technology, e-mail, 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.

4.2 History of Development of Digital Libraries

The historical significance of digital publishing is equivalent to that of
printed material and has significant effects on every aspect of the future of
library and information professionals. Digital libraries have transformed the
basic function of the library that of meeting the information needs of the
user. Paradigm change is being witnessed in the ways information is being
produced, delivered and accessed.

The origin of the digital library can be traced back to the writings of H G
Wells (1938), the well known English science fiction pioneer specially in
‘World Brain’ (Wells, 1938) a collection of essays where he introduced the
concept of World Encyclopedia. The concept of Memex in ‘As We May
Think’, an essay by Vannevar Bush, first published in The Atlantic Monthly, in July 1945 (Bush, 1945). Engelbart (1962), an early computer and Internet pioneer best known for his work on the challenges of human–computer interaction, reasoned that the state of our current technology controls our ability to manipulate information, and that fact in turn will control our ability to develop new, improved technologies. Thus, he set himself to the revolutionary task of developing computer-based technologies for processing information directly and also to improve individual and group processes for knowledge work. Engelbart’s philosophy and research agenda is most clearly and directly expressed in his research report. Licklider (1965) in his book ‘Libraries of the future’ (1965) puts his networked systems theories together for information service and retrieval and argued that libraries may be replaced by computers in the future. The contribution of the above stalwarts has a lot of impact on the present day digital libraries.

Providing access to information sources in response to user requests has seen the libraries transitioning from collecting material to developing policies for doing the same. This has resulted in the new phenomenon of ‘just in time’ from the age old ‘just in case’ to cater to the user’s needs. The shift to on-demand delivery can be facilitated by providing access to digital material from any source and any place. This digital transition has resulted in research being undertaken on various aspects of building digital library. So all sources where the information is available in electronic formats and accessible with the help of computers is referred in a wider perspective as
'electronic information resources' and encompass terms like automated library, electronic library, virtual library, networked library, multimedia library. All these are being used interchangeably and synonymously. But the term 'digital library' has become the preferred term due to the integration of mission, techniques and cultures of physical libraries with the abilities and culture of computing and telecommunications. Since digital information can be delivered directly to the user, and can be used simultaneously by many, is cost effective, has sophisticated searching and fast retrieval is well established and understood. The last two decades has witnessed the spurt of Digital Library Initiatives all over the world. Most of them are a result of the digital library research and development activities.

4.3 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)

According to E.A. Fox (Fox, 1999) the digital library may be defined as the “New way of carrying out the functions of libraries encompassing new types of information resources, new approaches to classification and cataloguing, intensive use of electronic systems and networks and dramatic shifts in intellectual, organizational and electronic practices”. The 1994 IEEE CALA Workshop on Intelligence Access to Online Digital Libraries gave a very comprehensive definition of a digital library which
A digital library is an assemblage of digital computing, storage and communications machinery together with the content and software needed to reproduce, emulate, and extend the services provided by conventional libraries based on paper and other material means of collecting, cataloging, finding, and disseminating information. A full service digital library must accomplish all essential services of traditional libraries and also exploit the well known advantages of digital storage, searching, and communication.

Borgman (1999) identifies two distinct senses in which "digital library" has been used (p. 227). The technological definition stating that "digital libraries are a set of electronic resources and associated technical capabilities for creating, searching and using information" (p. 234), is contrasted by the social view stating that "digital libraries are constructed, collected and organized, by (and for) a community of users, and their functional capabilities support the information needs and uses of that community" (p. 234). In 1999, Oppenheim and Smithson (Oppenheim & Smithson, 1999) defined digital library as "an information service in which all the information resources are available in computer process able form and the functions of acquisition, storage, preservation, retrieval, access and display are carried out through the use of digital technologies".

A digital library is a library in which collections are stored in digital formats (as opposed to print, microform, or other media) and accessible by computers. The DELOS Digital Library Reference Model
(http://www.delos.info ) defines a "digital library as an organization, which might be virtual, that comprehensively collects, manages and preserves for the long term rich digital content, and offers to its user communities specialized functionality on that content, of measurable quality and according to codified policies".

According to the Digital Library Federation (DLF, USA -http://www.dlf.org), "Digital Libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities".

To sum up, digital libraries are networked collections of digital texts, documents, images, sounds, data, software, and many more that are the core of today's web and tomorrow's universally accessible digital repositories of all human knowledge.

4.4 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.

4.4.1 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.

4.4.2 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.

4.4.3 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.

4.5 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;

4.5.1 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 digital
libraries. 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.
4.5.2 Information Resources Organisation

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
- 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.

4.5.3 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 centres.

- **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.

In selecting a metadata standard it is important that it should meet the requisite purpose. In other words, it may able to describe the content and meets the requirements such as access, preservation, management etc. It should also be transparent and preferably widely adopted to allow for easier interoperability, sharing and support. It is generally perceived that metadata is being different from the traditional cataloguing which requires cataloguing rules and stringent quality controls. Therefore, in the electronic media, there was a need of simpler light weight metadata which may be used for varying kind of organizations. Dublin Core is one such format having 15 elements to describe digital objects. One can follow the following four approaches available for the bibliographic control of electronic resources of varying control.  
(http://www.odi.ox.ac.uk\metadata.htm);
- Full MARC cataloguing for high quality resources that are likely to have continuing value.
- Enriched Dublin Core for next level.
- Minimal Dublin Core for Next level.
- Unstructured full-text keyword searching for the reminder.

**4.5.4 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 reorganisation, cession of sponsorship or support, technology obsolescence, content format obsolescence, hacking and sabotage of data and data files, disaster, whether nature or man-made 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.

4.5.5 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
In addition, several other services one can consider adding to digital libraries to promote the access to collection available.

4.6 Setting up Digital Libraries

While development of digital libraries it important to take in to consideration the following major aspects;

- Challenges
- General Principles
- Digital Collection Development

4.6.1 Challenges in Building Digital Libraries

Developing digital libraries being complex and long term activity in terms sustenance and providing uninterrupted access, one has to consider
Facing several challenges come in the way. Some of the major challenges that must be met to build effective digital libraries are:

- Identify and use most recent and improved technology for digitizing analogue materials.
- Design tools that facilitate the enhancement of cataloguing or descriptive information by incorporating the contributions of users.
- Establish protocols and standards to facilitate the assembly of distributed digital libraries.
- Address legal concerns associated with access, copying, and dissemination of physical and digital materials.
- Integrate access to both digital and physical materials.
- Develop approaches that can present heterogeneous resources in a coherent way.
- Provide more efficient and more flexible tools for transforming digital content to suit the needs of end-users.
- Develop economic models for the support of the Digital Library.

Giving a thought to these aspects beforehand will ease up the process and also any unknown problems.

4.6.2 General Principles for Digital Library Development

Some of the general principles to be followed in building digital libraries are,
• Technology, methods, materials and tools are changing and becoming obsolete swiftly in the area of Digital Library. Thus, in depth planning is required in order to provide services for a long time.

• Content is the crucial component of a Digital Library. One has to make wise decision to select the right content for the Digital Library.

• Involving computer and library professionals can help in offering expertise in building Digital Libraries. Along with it the patronage of Senior Management is essential.

• Lacking organisational commitment, it may not make sense to even begin a Digital Library project.

• Digital Libraries are mostly made available on Internet, therefore, one has to create a simple interface and provide multiple access points. It is should be mandate to consider accessibility for users with range of physical disabilities.

• One has to avoid proprietary hardware and software solutions whenever possible.

• In digitization projects, every attempt should be made to seek permission from the copyright owner for the materials that are to be digitized.

• The use of standards ensures more scalability, interoperability and portability.

• Technological obsolescence is the greatest threat to Digital Collections. "Migration" strategy needs to be planned. Persistence implies a commitment to both maintaining the object and keeping it accessible.
for a long time to come.


4.6.3 Digital Collection Development

The National Information Standards Organization (NISO 2007) after several rounds of discussions and consultations has come up with the following seven broad principles for developing digital collection. These principles serves as basis for developing any type digital library collection that one would like to create. These principles in brief are,

- Explicit collection development policy
- Scope, access restrictions and other characteristics to be described
- Should be sustainable
- Usable by a wide variety of people
- Respect intellectual property rights
- Mechanisms to record information demonstrating use of the collection
- Fitting into larger initiatives

(NISO http://www.niso.org/publications/rp/framework3.pdf)

Even if one is developing a digital library for school children, in this case, Children National Science Digital Library for India, these principles guidelines will have to be kept in mind.
4.7 Broad Phases in Developing Digital Collection

Developing digital library collection itself is a mega project. In order to successfully implement the project, it is important to follow the set of project management guidelines and techniques. Mark Jordan in his book *Putting Contents Online* suggested the following 10 phases to be followed for this purpose;

- Define goals and scope of the collection
- Evaluate and select source material to be added to collection
- Clear permission (copyright and intellectual property right issues) to use the source material
- Define project objectives and milestones
- Determine technical specifications
- Develop workflows
- Develop budget
- Execute the project
- Evaluate the project
- Evaluate the collection (Jordan 2006)

These phases will be equally applicable for creating any type of digital collection including the proposed in this research study.
4.8 Steps in Digitisation & Developing Digital Collection

As mentioned in the previous sections, digitisation is an important but most complex activity. One has to follow the set workflow systematically and step by step. Some of the steps involved in the digitisation process are listed and described below;

4.8.1 Selection of Materials

Content is the most important component of a Digital Library. Making the right decision regarding the content is a must for the Digital Library. One must ascertain what items to digitise if they exist in print form, metadata describing content and other attributes of each item.

4.8.2 Copyright Clearance

Most times, the materials that one selects for digitisation may not be owned by the one who is doing the digitisation. Therefore, it important to know who owns the copyright of each piece of material selected and then formally ask permission to digitise. Otherwise one has to check, ascertain and ensure that the selected item is out of copyright period. Under no circumstance, those involved in the process are expected to violate the copyright.
4.8.3 Digitization

Digital documents may be born-digital, created using digital publishing tools (e.g. Word, LaTex, DTP), or created by converting from an analogue format to digital format or converted from one digital format to another to suite the requirements of a particular Digital Library. The process of capturing and converting from analogue to digital format is often called as 'digitization' or 'digitalization'. It offers number of benefits as given below;

- Substantial Savings in administrative costs
- Secured medium and easy back-up options of your critical data
- Multiple user access to digital files
- Accessing information as a value added services by providing keywords search facilities and links to the resources such as full text searching, browsing, etc.
- Easy accessibility of data on your finger tips
- Disaster management in case of loss of physical data
- Zero wear and tear and deterioration
- Life time preservation of physical documents

4.8.4 Scanners and Scanning

The first step in converting paper documents into a digital library collection is to obtain images of all pages of all publications in digital format. Digital scanners are used to capture a digital image from an analogue media
such as printed page or a microfiche/microfilm at a predefined resolution and dynamic range (bit range). These are of two types:

- **Vector Scanners**
- **Raster Scanners**

**Vector Scanners**

It scans an image as a complex set of x,y coordinates. Vector images are generally used in geographical information system (GIS). The display software for the vector image interprets the image as function of coordinates and other included information to produce an electronic replica of the original drawing or a map. Maps, engineering drawings, and architectural blueprints are often scanned as vector images.

**Raster Scanners**

Raster images are captured by raster scanners by passing lights down the page and digitally encoded it row by row. The scanners used for digitizing analogue images into digital images come in a variety of shapes and sizes. Some of the common raster scanners are:

- Low-cost flat-bed scanner: Low-cost flat-bed units are the cheapest and most widely available type of scanner.
- Low-end scanner with sheet feeder:
- Professional duplex scanners: Professional scanners are reliable, heavy-duty machines capable of processing a large volume of pages.

- **OCR: Optical Character Recognition Software**

The following steps are involved in converting paper document into digital text format which is exactly replica of the paper document. It has the following steps:

- scanning;
- page layout analysis;
- reading;
- scanning tables and images marked as image.

- **Image Capturing Software**

Every scanner comes with its own software, which means that the program must be installed on the computer that manages the scanner. Some have a computer card that needs to be installed in the computer to speed up the scanning operation. Some of Image Capturing Software are given as follows:

- ABBYY Fine Reader (http://finereader.abbyy.com)
- Altris Software (http://www.altris.com/)
- OPTIX (http://www.blueridge.com/)
- Poweroffice (http://www.expower.com/)
- Read-Iris (http://www.readiris.com/)
- Omnipage (http://www.omnipage.com/)

- Document Preparation

A digital image of each page is scanned and transformed into a Bitmap or TIFF image using image capturing software provided with the scanner or using commercial scanning software.

4.8.5 Metadata Preparation and Tagging

Metadata is stored in TIFF files in its header whereas other metadata schemes such as Dublin Core and Metadata Encoding and Transmission Standard (METS) hold information about the digital object in a separate independent file. Keywords should be assigned to each digital object using standard vocabulary. Similarly, in case of authors, proper name authority file should be maintained so that name of persons may be assigned in a standardized form.

4.8.6 Naming convention of Files

File naming convention is very important in a digitization project as it helps in identifying the file for reference purpose. It should be applicable to the original TIFF files as well to all the derived files such as HTML, XML, Word, Pdf.
4.8.7 Selection of Software

It is important to have robust and flexible digital collections management and preservation software for creating and delivering digital collections. The complexity of the situation is that digital library technologies and contents are not static. Continual evolution and investment are required to maintain the digital library. One has to choose either of two, that is, commercial vs open source software.

Commercial digital library products are comprehensive and extensible enough to support this evolution, but are beyond the reach of most of the libraries in India due to:

- Initial purchase fee
- Licensing fee
- Upgrade fee
- Annual maintenance contacts (AMCs)

There is a long list of commercially available softwares for the purpose of creating digital collection, but prominent one which is widely used is CONTENT dm designed, developed, marketed and supported by OCLC.

The easiest choice for the librarian at present is to turn to Open Source Software (OSS). OSS has grown tremendously in scope and popularity over the last several years, and is now in widespread use. The growth of
OSS has gained the attention of research librarians and created new opportunities for libraries. OSS offers the following benefits:

- It is not privately controlled and hence likely to promote open rather than proprietary formats.
- It is typically maintained by communities rather than corporations and hence bug fixes and enhancement are often frequent and free.
- It is usually distributed free of charge

There are many softwares available which are popular among Indian libraries like Green Stone Digital Library, DSpace, Eprints, Fedora as Open Source Software. The criteria to choose the digital library software can be based on:

- Architecture of software
- Standards followed
- Interoperability
- Customisation
- License under which software issued
- Updating
- Regular users meet conducted
- Support
- Users base of software
4.8.8 Long Term Policy of Building Digital Library

It is essential that digital library project should sustain for a longer period of time. One needs to ensure that digital preservation planning is in place. It includes using standard file formats, migrating and refreshing files regularly, using an archival repository or other reliable offset storage, choosing appropriate storage media, and one has to see to it that adequate funding and organizational support are intact.

4.8.9 Migration, Retention Schedule and Refreshing

The technological tools are becoming obsolete very fast and so are the standards, formats, softwares and other tools used for developing digital collections. It is important to take in to account how to migrate to next generation platform without loss of any data and affecting access to users. Specialised agencies and professional bodies in this area have come up with possible solutions which one has to evaluate and adopt from time to time.

4.9 Characteristics of Digital Libraries

The characteristics of a typical digital library of any type or size can be described as follows;

- Text, images, audio and video are the various digital information resources contained in the digital libraries
- The physical space required to maintain traditional libraries and print resources, is reduced due to digital libraries
- Multi level services are developed to meet the need of local as well as remote users as it’s a well known fact that digital libraries may be situated in any part of the world
- Infrastructure, interoperability are important factors in digital library development and management as digital libraries provide access to varied information resources located on different servers around the world
- Digital libraries enable many users to use the same information source at the same time
- Digital libraries usually provide access to materials that one doesn’t own of which some are available for free and some may be provided upon payment of a fee.
- Digital libraries use appropriate mechanisms for filtering out what is unwanted as the problem faced is that of information overload.
- Digital libraries are able to handle multilingual information resources.
- Appropriate mechanisms are put in place to support users with all the different levels of IT, subject and linguistic skills.
- Digital libraries allow enhanced searching and retrieval facilities
- Different users can view and use digital information to suit their individual needs.
- Users from any part of the world are able to use digital library at any
time and in any language thus breaking the barrier of time, space
and language. (Chowdhury, 2003).

4.10 Advantages of Digital Libraries

Libraries in physical form are in existence since several centuries and
have reformed the way the society works. But it is incomparable when it
comes to the influence and impact the present day digital libraries have
exerted on the societies. The knowledge society is reaping the benefits of
the digital library in day to day activities which is many folds compared to
the benefits of a traditional library. The information revolution not only
supplies the technological horsepower that drives digital libraries but also
fuels an unprecedented demand for storing, organizing and accessing
information. 'If information is the currency of the knowledge economy,
digital libraries are the banks where it is invested'. (Ian H. W. and others,
2010)

Some advantages of a digital library over a conventional library are,
- cutting down costs of library maintenance
- resource distribution
- resources are electronically catalogued and can be browsed
- provides equitable and widely distributed access at lower costs
- most appropriate means of organizing intellectual artifacts that
cannot be represented or distributed in printed formats,
- evolves into a complex system that makes information available in hard copy, on magnetic tape/discs, CD-ROMS, including those from online sources.

Important applications identified for digital libraries are,

- Archival preservation such as manuscripts, ancient literary works, cultural artifacts, community identities;
- Legal documentation such as government documentation of plans and policies, history of legal cases, census and statistical data, spatial data and other relevant information that can be brought into the public domain; and
- Educational and research purposes such as scholarly publishing, theses, research work, hosting reference material.

4.11 Digitization Initiatives in India

Digital libraries in India began in the mid-1990s in comparison to the developed countries where they started during the 1970s and it was largely due to availability of information technology (IT) on a large scale and the support extended by the central government. The advent of the Internet acted as a catalyst for digital library initiatives. Digital library projects and developments in the country are many, and a large number of them are at the preliminary stage.
• The Indian Institute of Science (IISc.), Bangalore, collaborated with IBM and sun Microsystems to host, respectively, the IBM digital library and SunSite in their campus.

• The Indian Academy of Science, Bangalore, has demonstrated successfully free on-web delivery of their journals through their website.

• The Council of Scientific and Industrial Research (CSIR) of India has created a major project to map the traditional knowledge of the country with the Traditional Knowledge Digital Library (TKDL) (http://203.200.90.6/tkdl/langdefault/common/home.asp).

• The Central University of Hyderabad in collaboration with Sun has digitized some of its collection.

• The central library of IIT, Kharagpur, has been awarded a project by ARDB for developing a hypermedia digital library on aerospace science and technology.

• There is also a strong component of digital library in The NPTEL that focuses on the role of technology in knowledge accumulation, storing and dissemination and education in the three sectors of university, industry and government. NPTEL offers collection of hundreds of video lectures. This has been set up by the Indian Institute of Technologies (IITs), Indian Institute Of Managements
(IIMs) and Carnegie Mellon University, aimed at providing distance education, developing resources for core courses, conducting joint Ph.D. Programmes and setting up a digital library.

• Training on Greenstone digital library software has been conducted by NCSI of IISc, Bangalore. Inspired by the success of NDTLD, University of Mysore is pursuing a project, Vidyanidhi, (http://www.vidyanidhi.org.in) on digitizing theses and dissertations.

• Shodhganga (http://shodhganga.inflibnet.ac.in/), a major initiative of INFLIBNET Centre supported by UGC/MHRD is the digital reservoir of Indian theses and dissertations. This digital library hosts full text of theses submitted to Indian Universities.

A few other noteworthy initiatives on Digital Libraries in India in various areas are given below:

• Million Book Universal Digital library Project - Carnegie Mellon - IISc – ERNET (http://www.dli.ernet.in) The Digital Library of India (DLI) aims to preserve Indian heritage that is available in the books, manuscripts, art and music. More than twenty “Content Creation Centres” are partners in the DLI. Major among them are Indian Institute of Science (IISc), Carnegie Mellon University (CMU), the International Institute of Information Technology Hyderabad (IIITH), etc.
• **Khuda Baksh Oriental Library** ([http://kblibrary.nic.in/](http://kblibrary.nic.in/))

This library has a rich collection of manuscripts in Persian, Arabic, Urdu and other languages. What was undertaken as digitisation project has resulted in the print catalogues being converted to machine-readable form by NICNET. Around 400 thousand pages are available and are accessible as JEG files.

• **Indira Gandhi Centre for the ARTS (IGNCA Digital Library)** ([http://ignca.nic.in](http://ignca.nic.in))

Sponsored by the Ministry of Communication and Information Technology (MCIT) IGNA has taken up the Kalasampada Digital Library – Resourcee for Indian Cultural Heritage (DL-RICH) project. The objective of this project is to develop software will allows users to interact and explore images, audio, text, graphics, animation and video to aid the study of Indian art and culture.

• **Nalanda Digital Library**, National Institute of Technology (NIT) Calicut ([http://www.nalanda.nitc.ac.in](http://www.nalanda.nitc.ac.in)) It is one of the largest digital libraries in the country and was initiated in the year 1999. It caters to the needs the research and information needs in the fields of science, engineering and technology. It contains theses, dissertations, course materials, articles and annual reports and uses a software designed by the institute itself.
• Archives of Indian Labour at the V.V. Giri National Labour Institute and the Association of Indian Labour Historians (AILH) (http://www.indialabourarchives.org/sources/jnu.htm)
The archives aim at preserving documents, builds collections and initiates research in labour history. It uses Greenstone – an open source digital library system to integrate text, audio and video.

• Librarian's Digital Library at DRTC Bangalore (https://drtc.isibang.ac.in/index.jsp)
It’s a repository at the DRTC, ISI, Bangalore. It uses Dspace and contains articles, theses and dissertations, presentations, multi-lingual documents, photographs, etc.

• Digitization of the official debates of Rajya Sabha (http://rsde debates.nic.in) collection of all the debates took place in Rajya Sabha.

These are only few examples of digital library initiatives which are listed to demonstrate the fact that, India is not behind any of the developing countries in the world. In the recent past, every academic institution, professional bodies and major government departments have taken steps to convert their print publications into digital collection. Many are going to be seen in the forthcoming year. All this is due to availability of robust technologies at reasonable cost, digital library softwares, good number of training programs and support from parent bodies.
4.12 Digital Library Initiatives for Children

A quick glance at the published literature reveals that, quite a few digital libraries have come up during past few years. Most of them being much localized, only two which are quite well known and are frequently used are described here.

4.12.1 National Science Digital Library (NSDL)

National Science Digital Library is an online portal concentrating on open educational sources in digital form and was formally launched in the year 2002. Basically it funded by National Science Foundation and caters to Science, Technology, Engineering and Mathematics (STEM) education. This digital library gives access to more than a million selected online STEM resources. The stress is on high quality, online educational resources for teaching and learning. A collaborative team of principal investigators and staff consisting of researchers, content providers, developers, librarians, teachers and students manage NSDL. A core integration team work with 9 ‘pathways’ portals and 200 NSF grantees. Most resources follow the principles of Open Educational Resource (OER).

The NSDL network provide digital collections, services, partnering opportunities, etc over various educational levels like K-12 (Kindergarten to class 12), higher education and disciplines. NSDL has a simplified user interface and thus is easy to use. NSDL makes searches using NSDL metadata records and the text of the resources if the text is available
publicly on the Internet. NSDL includes images, video, audio, animations, software, games, lesson plans, tutorials, activities, etc and provides a user help and an array of technical tools and services, metadata and collection creation and management services along with search and browse facility. (https://nsdl.org/)

4.12.2 International Children’s Digital Library (ICDL)

National Science Foundation, Institute of Museum and Library Services and Microsoft Research have funded ‘The International Children’s Digital Library (ICDL)’ as a research project in November 2002. Helping children understand the world around them through the resources in the original languages is the main aim of the collection. ICDL collection is expected to accumulate around 10,000 resources in minimum 100 languages.

An interdisciplinary research team consisting of computer scientists, librarians, educational technologists, classroom teachers, children, etc. at the University of Maryland in cooperation with the Internet Archive created the ICDL. Involvement of children and adults called as cooperative inquiry was engaged to develop and evaluate computer interface technologies that support searching, browsing, reading and sharing books in electronic form.

ICDL special feature is that ‘Born Digital’ books are not included in the collection. Only physical books that have been published and converted into digital format are added. (Weeks Ann Carlson and others, 2003)
4.13 Current Trends in Digital Library

Library and Information science being a discipline has had the trend of adopting innovative practices which aids in effectiveness and efficiency of the services. Transcending from semi digital libraries to total digital libraries will be inevitable in the 21st century. Some of the emerging trends that will impact libraries and librarianship in a big way are presented here.

Total multimedia solution will be accessible to a large number of institutions as there will be improvement in streaming technologies, and network bandwidth will be more economical. Accepted standards will be incorporated by software vendors as they eye the library market which is at present in the discussion phase and the experimental stage of trying the standards. A successful trend would be the transition of digital libraries from being computer science experiments to being major library implementations. (Sujin, 2013)

Many faculty members of leading universities are managing their teaching through course specific web pages like Blackboard and WebCT wherein necessary support facilities like online discussion forums, automatic grading, etc are provided. But the concern as expressed by David Seaman in his paper ‘Aggregation, Integration, and Openness: Current Trends in Digital Libraries, such courseware systems are often installed and run by IT departments without much library involvement and there is little human and technical interface between the library content management systems and the courseware systems. Also there is difficulty in linking the webpage course contents to the holdings of the digital library supporting it. He has
further observed that ‘the advent of courseware holds great promise for moving the library into the classroom but so far the reality falls short of this opportunity to engage even more richly with the work of our teachers and students’. (Seaman, David DLF.USA http://www.dlf.org) So this trend prompts us to have more of library involvement in course design and delivery for effective output.

Libraries being custodians of scholarly works have always focused on preservation and archiving and in the context of digital library too these two areas are of great future. In the developed countries institutional repositories are being created to store the intellectual assets of the scholars in the form of databases, images, teaching modules, etc so that it will be a valuable material for others to reuse as the new generation of researchers are already in the digital landscape. Some activities worth mentioning in the area of digital preservation is Digital Curation Centre to provide national research and advice on the storage and lifecycle management of digital objects funded by JISC, The National Digital Information Infrastructure and Preservation Program led by the Library of Congress, etc. (David Seaman). These noteworthy activities can be taken as guiding projects and the digital preservation initiatives in the Indian academic institutes can be modeled.

4.14 Conclusion

Emerging information and communication technologies have impacted in big way which is difficult to measure in a quantifiable method. These
technological changes have forced publishing, editing, translating, database creation, computing, teaching, and learning to integrate to enhance the quality, speed and coverage in various fields. A large part of the world's information is now produced in digital form. It is a known fact that World Wide Web has become the basic infrastructure of everyday life. Everyday Millions of people all over the world search the Web for one purpose or the other. For transformation to happen, collections of all kinds must be digitized and indexed effectively, from small communities to large disciplines, from formal to informal communications, from text to image and video repositories, and eventually across languages and cultures. Digital collection in various formats forms the basis of digital library. It is interesting to note that, a digital library may achieve the status of what Dr. S R Ranganathan called a “Library is a Growing Organism”. A positive expectation would be that through its activities, a digital library may become a “learning organism”. Digital libraries are a result of the evolution of libraries for centuries. They represent a giant leap forward in information services for better collaboration and productive output. The future is unpredictable, yet there are instances of intuitions. Visionaries like Vannevar Bush and Ted Nelson who have dreamt of “linking all the books in the world”, seemed impossible in 1970’s, but digitization has proved that it can be a reality. It is natural for libraries to provide equitable access to digital data stream for their users. It is the need of the hour for libraries to choose right digital platform to suit their users and their institutional goals.
Digital Libraries have evolved phase wise and in a systematic manner throughout the world. Yet they have presented new challenges to library and information professionals. Having well established digital libraries in the developed countries has helped in guiding initiatives in India. Guidelines regarding the standards, technologies, dynamics are very precise and can act as a compass to new digital library initiatives. As library and information professionals are well aware of the advantages of digital libraries over traditional libraries it is a matter of proactive decision to design and develop digital libraries to suit their parent organizations and their users for providing effective and efficient services. The concern is how to develop the digital libraries in a cost effective way and in an easy manner.

REFERENCES


