DIGITAL LIBRARY

INTRODUCTION

A **digital library** is a special library with a collection of **digital objects** that can include **text**, **visual material**, **audio material**, **video material**, stored as **electronic media formats** (as opposed to print, or other media), along with means for **organizing, storing, and retrieving the files and media** contained in the library collection. Digital libraries can **vary immensely** in **size and scope**, and can be maintained by **individuals, organizations, or affiliated** with established physical library buildings or institutions, or with academic institutions. The digital content may be stored **locally**, or **accessed remotely via computer networks**. An electronic library is a type of information retrieval system. These information retrieval systems are able to exchange information with each other through interoperability and sustainability.

DEFINITION OF TERMS

DIGITAL LIBRARIES -   was first popularized by the NSF/DARPA/NASA Digital Libraries Initiative in 1994. These draw heavily on Vannevar Bush's essay As We May Think (1945), which set out a vision not in terms of technology, but user experience.

-       a potentially virtual organization, that comprehensively collects, manages and preserves for the long depth of time rich digital content, and offers to its targeted user communities specialized functionality on that content, of defined quality and according to comprehensive codified policies (DELOS, a Network of Excellence on Digital Libraries, and DL.org, a Coordination Action on Digital Library Interoperability, Best Practices and Modelling Foundations, Digital Library researchers and practitioners and software developer produced a Digital Library Reference Model

-       A distinction is often made between content that was created in a digital format, known as born-digital, and information that has been converted from a physical medium, e.g. paper, through digitization.

VIRTUAL LIBRARY   -   was initially used interchangeably with digital library, but is now primarily used for libraries that are virtual in other senses (such as libraries which aggregate distributed content). In the early days of digital libraries, there was discussion of the similarities and differences among the terms digital, virtual, and electronic.

HYBRID LIBRARY   -   is sometimes used for libraries that have both physical collections and electronic collections. For example, American Memory is a digital library within the Library of Congress.

Software implementation

Institutional repository software is designed for archiving, organizing, and searching a library's content. Popular open-source solutions include **DSpace, EPrints, Digital Commons, and Fedora Commons-based systems Islandora and Hydra**. The design and implementation in digital libraries are constructed so computer systems and software can make use of the information

when it is exchanged. These are referred to as semantic digital libraries. Semantic libraries are also used to socialize with different communities from a mass of social networks. **DjDL** is a type of **semantic digital library**. **Keywords-based** and **semantic search** are the two main types of searches. A tool is provided in the semantic search that create a group for augmentation and refinement for keywords-based search. Conceptual knowledge used in DjDL is centered around two forms; the **subject ontology** and the **set of concept search patterns based on the ontology**. The three type of ontologies that are associated to this search are bibliographic ontologies, community-aware ontologies, and subject ontologies.

History

The concept of digital libraries **emerged** in **1892** from the early ideas of **Paul Outlet** in ways to **cease the violent wars, eliminate national boundaries, and allow humanity to become balanced**. He discussed in his book called "**Birth of the Information Age**" about how to **interlink millions** of **documents, images, audio and video files** together so **people could search in one system**. He called it the "**Mundaneum**." In present time, this idea is closely associated with the **Internet**. **Vannevar Bush** and **J.C.R. Licklider** are two more contributors that advanced this idea into newer technology. Bush was seen as a researcher that assisted in making 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 individuals 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 **procognitive** 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) which was "born digital" in 1966.

In 1994, digital libraries became popular due to the $24.4 million that the U.S. federal funded among six universities for research. The universities included Carnegie Mellon University, University of California-Berkeley, University of Michigan, University of Illinois, University of California-Santa Barbara, and the Stanford University.

Academic repositories

Many academic libraries are actively involved in building institutional repositories of the institution's books, papers, theses, and other works which can be digitized or were 'born digital'. Many of these repositories are made available to the general public with few restrictions, in accordance with the goals of open access, in contrast to the publication of research in commercial journals, where the publishers often limit access rights. Institutional, truly free, and corporate repositories are sometimes referred to as digital libraries.

Some important digital libraries also serve as long term archives, such as arXiv and the Internet Archive. Others, such as the Digital Public Library of America, seek to make digital information from various institutions widely accessible online.

Digital archives

Physical archives differ from physical libraries in several ways. Traditionally, archives are defined as:

Containing primary sources of information (typically letters and papers directly produced by an individual or organization) rather than the secondary sources found in a library (books, periodicals, etc.).

- Having their contents organized in groups rather than individual items.
- Having unique contents.

The technology used to create digital libraries is even more revolutionary for archives since it breaks down the second and third of these general rules. In other words, "digital archives" or "online archives" will still generally contain primary sources, but they are likely to be described individually rather than (or in addition to) in groups or collections. Further, because they are digital, their contents are easily reproducible and may indeed have been reproduced from elsewhere. The Oxford Text Archive is generally considered to be the oldest digital archive of academic physical primary source materials.

Archives differ from libraries in the nature of the materials held. Libraries collect individual published books and serials, or bounded sets of individual items. The books and journals held by libraries are not unique, since multiple copies exist and any given copy will generally prove as satisfactory as any other copy. The material in archives and manuscript libraries are "the unique records of corporate bodies and the papers of individuals and families".

A fundamental characteristic of archives is that they have to keep the context in which their records have been created and the network of relationships between them in order to preserve their informative content and provide understandable and useful information over time. The fundamental characteristic of archives resides in their hierarchical organization expressing the context by means of the archival bond. Archival descriptions are the fundamental means to describe, understand, retrieve and access archival material. At the digital level, archival descriptions are usually encoded by means of the Encoded Archival Description XML format. The EAD is a standardized electronic

representation of archival description which makes it possible to provide union access to detailed archival descriptions and resources in repositories distributed throughout the world.

Digital libraries benefit from the existence of sophisticated formal models, such as The 5S Framework: Streams, Structures, Spaces, Scenarios and Societies, which allow us to formally describe them and to prove their properties and features.

Given the importance of archives, a dedicated formal model, called NEsted SeTs for Object Hierarchies (NESTOR), built around their peculiar constituents, has been defined. NESTOR is based on the idea of expressing the hierarchical relationships between objects through the inclusion property between sets, in contrast to the binary relation between nodes exploited by the tree. NESTOR has been used to formally extend the 5S model to define a digital archive as a specific case of digital library able to take into consideration the peculiar features of archives.

Searching

Most digital libraries provide a search interface which allows resources to be found. These resources are typically deep web (or invisible web) resources since they frequently cannot be located by search engine crawlers. Some digital libraries create special pages or sitemaps to allow search engines to find all their resources. Digital libraries frequently use the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) to expose their metadata to other digital libraries, and search engines like Google Scholar, Yahoo! and Scirus can also use OAI-PMH to find these deep web resources.

There are two general strategies for searching a federation of digital libraries: distributed searching and searching previously harvested metadata.

Distributed searching typically involves a client sending multiple search requests in parallel to a number of servers in the federation. The results are gathered, duplicates are eliminated or clustered, and the remaining items are sorted and presented back to the client. Protocols like Z39.50 are frequently used in distributed searching. A benefit to this approach is that the resource-intensive tasks of indexing and storage are left to the respective servers in the federation. A drawback to this approach is that the search mechanism is limited by the different indexing and ranking capabilities of each database; therefore, making it difficult to assemble a combined result consisting of the most relevant found items.

Searching over previously harvested metadata involves searching a locally stored index of information that has previously been collected from the libraries in the federation. When a search is performed, the search mechanism does not need to make connections with the digital libraries it is searching - it already has a local representation of the information. This approach requires the creation of an indexing and harvesting mechanism which operates regularly, connecting to all the digital libraries and querying the whole collection in order to discover new and updated resources. OAI-PMH is frequently used by digital libraries for allowing metadata to be harvested. A benefit to this approach is that the search mechanism has full control over indexing and ranking algorithms,

possibly allowing more consistent results. A drawback is that harvesting and indexing systems are more resource-intensive and therefore expensive.

Software

There are a number of software packages for use in general digital libraries, for notable ones see Digital library software. Institutional repository software, which focuses primarily on ingest, preservation and access of locally produced documents, particularly locally produced academic outputs, can be found in Institutional repository software. This software may be proprietary, as is the case with the Library of Congress which uses Digiboard and CTS to manage digital content.

Digitization

In the past few years, procedures for digitizing books at high speed and comparatively low cost have improved considerably with the result that it is now possible to digitize millions of books per year. Google book-scanning project is also working with libraries to offer digitize books pushing forward on the digitize book realm.

Recommendation systems

Many digital libraries offer recommender systems to reduce information overload and help their users discovering relevant literature. Some examples of digital libraries offering recommender systems are IEEE Xplore, Europeana, and GESIS Sowiport. The recommender systems work mostly based on content-based filtering but also other approaches are used such as collaborative filtering and citation-based recommendations.[22] Beel et al. report that there are more than 90 different recommendation approaches for digital libraries, presented in more than 200 research articles.

Typically, digital libraries develop and maintain their own recommender systems based on existing search and recommendation frameworks such as Apache Lucene or Apache Mahout. However, there are also some recommendation-as-a-service provider specializing in offering a recommender system for digital libraries as a service.

Advantages

The advantages of digital libraries as a means of easily and rapidly accessing books, archives and images of various types are now widely recognized by commercial interests and public bodies alike.

Traditional libraries are limited by storage space; digital libraries have the potential to store much more information, simply because digital information requires very little physical space to contain it. As such, the cost of maintaining a digital library can be much lower than that of a traditional library. A physical library must spend large sums of money paying for staff, book maintenance, rent, and additional books. Digital libraries may reduce or, in some instances, do away with these fees. Both types of library require cataloging input to allow users to locate and retrieve material. Digital libraries may be more willing to adopt innovations in technology providing

users with improvements in electronic and audio book technology as well as presenting new forms of communication such as wikis and blogs; conventional libraries may consider that providing online access to their OP AC catalog is sufficient. An important advantage to digital conversion is increased accessibility to users. They also increase availability to individuals who may not be traditional patrons of a library, due to geographic location or organizational affiliation.

No physical boundary. The user of a digital library need not to go to the library physically; people from all over the world can gain access to the same information, as long as an Internet connection is available.

Round the clock availability. A major advantage of digital libraries is that people can gain access 24/7 to the information.

Multiple access. The same resources can be used simultaneously by a number of institutions and patrons. This may not be the case for copyrighted material: a library may have a license for "lending out" only one copy at a time; this is achieved with a system of digital rights management where a resource can become inaccessible after expiration of the lending period or after the lender chooses to make it inaccessible (equivalent to returning the resource).

Information retrieval. The user is able to use any search term (word, phrase, title, name, and subject) to search the entire collection. Digital libraries can provide very user-friendly interfaces, giving click able access to its resources.

Preservation and conservation. Digitization is not a long-term preservation solution for physical collections, but does succeed in providing access copies for materials that would otherwise fall to degradation from repeated use. Digitized collections and born-digital objects pose many preservation and conservation concerns that analog materials do not. Please see the following "Problems" section of this page for examples.

Space. Whereas traditional libraries are limited by storage space, digital libraries have the potential to store much more information, simply because digital information requires very little physical space to contain them and media storage technologies are more affordable than ever before.

Added value. Certain characteristics of objects, primarily the quality of images, may be improved. Digitization can enhance legibility and remove visible flaws such as stains and discoloration.

Easily accessible.

Digital preservation

Digital preservation aims to ensure that digital media and information systems are still interpretable into the indefinite future. Each necessary component of this must be migrated, preserved or emulated. Typically lower levels of systems (floppy disks for example) are emulated, bit-streams (the actual files stored in the disks) are preserved and operating systems are emulated

as a virtual machine. Only where the meaning and content of digital media and information systems are well understood is migration possible, as is the case for office documents. However, at least one organization, the Wider Net Project, has created an offline digital library, the eGranary, by reproducing materials on a 6 TB hard drive. Instead of a bit-stream environment, the digital library contains a built-in proxy server and search engine so the digital materials can be accessed using an Internet browser. Also, the materials are not preserved for the future. The eGranary is intended for use in places or situations where Internet connectivity is very slow, non-existent, unreliable, unsuitable or too expensive.

Copyright and licensing

Digital libraries are hampered by copyright law because, unlike with traditional printed works, the laws of digital copyright are still being formed. The republication of material on the web by libraries may require permission from rights holders, and there is a conflict of interest between libraries and the publishers who may wish to create online versions of their acquired content for commercial purposes. In 2010, it was estimated that twenty-three percent of books in existence were created before 1923 and thus out of copyright. Of those printed after this date, only five percent were still in print as of 2010. Thus, approximately seventy-two percent of books were not available to the public.

There is a dilution of responsibility that occurs as a result of the distributed nature of digital resources. Complex intellectual property matters may become involved since digital material is not always owned by a library. The content is, in many cases, public domain or self-generated content only. Some digital libraries, such as Project Gutenberg, work to digitize out-of-copyright works and make them freely available to the public. An estimate of the number of distinct books still existent in library catalogues from 2000 BC to 1960, has been made.

The Fair Use Provisions (17 USC § 107) under the Copyright Act of 1976 provide specific guidelines under which circumstances libraries are allowed to copy digital resources. Four factors that constitute fair use are "**Purpose of the use**, **Nature of the work**, **Amount or substantiality used** and **Market impact**."

Some digital libraries acquire a license to lend their resources. This may involve the restriction of lending out only one copy at a time for each license, and applying a system of digital rights management for this purpose.

The **Digital Millennium Copyright Act of 1998** was an act created in the United States to attempt to deal with the **introduction of digital works**. This Act incorporates two treaties from the year 1996. It **criminalizes** the attempt to **circumvent measures** which limit access to **copyrighted materials**. It also criminalizes the act of attempting to **circumvent access control**. This act provides an exemption for **nonprofit libraries** and **archives** which allows up to **three copies to be made**, **one of which may be digital**. This may not be made public or distributed on the web, however. Further, it allows libraries and archives to copy a work if its format becomes obsolete.

Copyright issues persist. As such, proposals have been put forward suggesting that **digital libraries be exempt from copyright law**. Although this would be very **beneficial** to the **public**, it may have a **negative economic effect** and **authors** may be **less inclined** to create **new works.**

Another issue that complicates matters is the desire of some **publishing houses** to **restrict** the use of digit materials such as **e-books purchased by libraries**. Whereas with **printed books**, the library **owns the book** until it can **no longer be circulated**, publishers want to **limit the number of times an e-book** can be checked out before the library would need to repurchase that book. "[HarperCollins] **began licensing use of each e-book copy for a maximum of 26 loans**. This affects only the most popular titles and has no practical effect on others. After the limit is reached, the library can repurchase access rights at a lower cost than the original price." While from a publishing perspective, this sounds like a good balance of library lending and protecting themselves from a feared decrease in book sales, libraries are not set up to monitor their collections as such. They acknowledge the increased demand of digital materials available to patrons and the desire of a digital library to become expanded to include best sellers, but publisher licensing may hinder the process...

Metadata creation

In **traditional libraries**, the **ability to find works of interest** is **directly related** to **how well** they were **cataloged.** While **cataloging electronic works digitized** from a library's existing holding may be as simple as copying or moving a record from the print to the electronic form, complex and born-digital works require substantially more effort. To handle the growing volume of electronic publications, new tools and technologies have to be designed to allow effective automated semantic classification and searching. While full text search can be used for some items, there are many common catalog searches which cannot be performed using full text, including:

- finding texts which are translations of other texts
- differentiating between editions/volumes of a text/periodical
- inconsistent descriptors(especially subject headings)
- missing, deficient or poor quality taxonomy practices
- linking texts published under pseudonyms to the real authors (Samuel Clemens and Mark Twain, for example)
- differentiating non-fiction from parody (The Onion from The New York Times)

Disadvantages

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.

- Bibliographic database
- Digital Library Federation
- D-Lib, a magazine dedicated to digital library research and development
- Digital humanities
- Full text database
- Mobile library
- Online encyclopedia
- Traveling library

Digital Libraries as contents collected on behalf of user communities, while librarians focused on digital libraries as instructions or services many new initiatives came through the research sponsored by the US National Science Foundation (NSF) and U.K. Joint information system committee (JISC). During the past decade thousands of digital libraries in variety of forms were built globally and are functioning operationally, with more to come. Hundreds of research.

A fully developed digital library environment involves the following elements.

1. Storage of digital content and metadata in an appropriate multimedia repository.
2. The extraction or creation of metadata to assist in object viewing. Management and Presentation.
3. Initial conversion of content from physical to digital form.
4. Patron access through a browser or dedicated client.
5. Content delivery via file transfer or streaming media.
6. Client services for the browser, including repository querying and workflow.

Why Digital Library

1. The Digital library bring the library to the users
2. Information can be shared
3. Information is always available

4. Computer power is used for searching, and browsing
5. It is easier to keep information current

Features of Digital Libraries

Digital library is a special library with a focused collection of digital objects that can include text, visual material, audio material, video material, stored as electronic media formats. Digital libraries possess huge collection of digital information having high value, high material quality, catalogued material, accurate origin, date and other external deceptive information.

Different features of any digital library is that it provide different services such as browsing & searching, Translation, Efficient Retrieval Functionality, Altering services, High services, High performance document servers, Document delivery, Data Dissemination, Intelligent user agents, Classification, user access model, publishing, etc.

Digital libraries possess certain qualities like they provide up-to-date, timely, high quality multimedia, which removes physical and conceptual barriers. They create rich virtual space by network connectivity and interactive technologies. They provide advance and innovative services.

They Provide following common browse and retrieval features like field specific search, proximity like field specific search, wild card, limit field searches, range searching , Boolean operators, Phrase specific, subject search and steaming, density of term, frequency of terms reference links, Persistent links , etc. On this basis user can search options and can conduct search on various fields.

Future trends in Digital Libraries

Exiting digital repository services are in an evolving stage, according to a survey there are more than 900 repositories open access running. There are many open source software which are helpful in creating open access repositories. As a result of which many more libraries will join in creating institutional repositories or digital libraries or digital repositories. So, there will be many digital repositories in near future.

Open access to information, vast content collection will be available with open source software tools and new method will be initiated of digital preservation. There is various software which are used for creating digital libraries, used by users. It is important to know the requirements of different class of user, and for that evaluation of various software's used for creating digital libraries is necessary. The software should be studied on the basis of that for which user community it is developed, how it is useful to users, how it can be improved to further help users to retrieve

information effectively. Not only user based evaluation but other evaluation of digital library system on the basis of their performance, information retrieval techniques & underlying technology used. The most difficult issue in digital library is evaluation which is to be concluded from the above stated points.

With increasing technology in the world, it is also important for librarians to learn and adopt new technologies and help the end users to use and make them available their valuable collection by the means of internet through digital library. As the technology keeps on improving new technologies comes into existence, leading to technical and operational challenges for libraries, digital library, archives and museums. For which they have to migrate to new technologies now and then.

So, librarians should learn and adopt new changes in the technology and understand their feature so that best software can be selected from all the available software. This will lead to improve performance and effectiveness of all those repositories in future and many new features can be added to the software.

Academic libraries are not only providing a broad array of electronic services to their primary clientele, but are also increasingly providing these services to off-campus users other than their primary clientele.

The past two decades have brought unprecedented changes in technology for academic libraries. Libraries once focused on helping users identify, retrieve, and use materials within the library building. Now they are supporting these activities with a broad array of electronic services, and increasingly, users can access these services from locations outside the library

Many academic libraries are taking advantage of technology to serve a broader clientele. Academic libraries at institutions with graduate programs have generally led in providing electronic services, but gaps between institution types are narrowing.

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.

According to Larry Lannom, Director of Information Management Technology at the nonprofit Corporation for National Research Initiatives (CNRI), "all the problems associated with digital libraries are wrapped up in archiving." He goes on to state, "If in 100 years people can still read your article, we'll have solved the problem." Daniel Akst, author of The Webster Chronicle, proposes that "the future of libraries — and of information — is digital." Peter Lyman and Hal Variant, information

scientists at the University of California, Berkeley, estimate that "the world's total yearly production of print, film, optical, and magnetic content would require roughly 1.5 billion gigabytes of storage." Therefore, they believe that "soon it will be technologically possible for an average person to access virtually all recorded information."

CONCLUSION

CHANGE is inevitable in every aspect of library work. Access to information is the core of the activities of libraries. The upsurge of digital sources of information and insurmountable development of telecommunications technology allow people to expand their information horizons. Because of these changes, libraries are behooved to move forward and take center stage. Through careful planning and working knowledge of information technology, librarians can face these challenges to ensure that the information resources and services that they provide meet the individual customer's varied needs.

REFERENCES

(Online Journal Article)

College resources: electronic services in academic libraries. (2005), Report by: U.S. Department of Education, U.S. Department of Education, p. 83. General OneFile, http://link.galegroup.com/apps/doc/A153701420/GPS?u=pup&sid=GPS&xid=2659f8b3. Accessed 11 Jan. 2018.

Shah, C. R. (2016). *Future of digital library*, Vol. *4* Issue No.5. http://www.raijmr.com/wp-content/uploads/2017/11/IJRHS_2016_vol04_issue_05_08.pdf. Accessed 10 Jan. 2018

(Online Article)

Digital library. (2017). https://en.wikipedia.org/wiki/Digital_library. Accessed 10 Jan. 2018

(Book)

Buenrostro, J. C. Jr. (2009). *The Digital Library and other issues in library and information science*. Quezon City : Great Books Publishing.