Digital Libraries in the New Millennium

Md. Zillur Rahman[1]

Abstract

The concept of digital library is not a new one, although it is only recently that digital library is gaining prominence all over the world. This article provides a holistic understanding of digital libraries including its present context and future directions. Various technical issues ingrained in the concept of digital library also has been analysed in the article. Issues like interoperability, information retrieval, property and access control and usability also has been discussed in detail. Although digital library is a technical phenomenon, it has wider social and economic implications. The article casts a critical glance at these implications as well. In the end, ongoing international efforts in the arena of digital libraries also has been discussed.

Keywords *Digital Library, New Technology, Interoperability, Storage, Information retrieval*

Introduction

From engraving on stone slates or writing on velum or Bhojpatra or skin, print culture outplaced the age and the paper which dominated for centuries in communication media, including library service. Now the electronic media has revolutionized the whole gamut to a new brand of information recording, multicopying and processing, including the Information Superhighway or WWW through different networking systems including Internet. The libraries are now information data banks, knowledge bases, text base or Gopherspace. We are moving towards a paperless society for variety of reasons. Right from morning newspaper to late night light reading before going to sleep, we are very much dependent on electronic media. In library and information profession, the new Information Technology provides a platform to achieve the targets with full precision, covering wide spectrum in the shortest possible time and in an economical way to reach any corner and user of the world. There is no discrimination of any kind and a free flow of information is ensured at the choice of the user.

Existing information carriers on paper or celluloid (books, journals, microforms; tapes, optical devices etc. and their surrogates like secondary or tertiary documents) are being converted into computer readable databases gradually to meet the present need of the society. The

[1]*Assistant Librarian, Independent University, Bangladesh. E-mail: zillurmd@gmail.com*

participating information generating agencies are now creating their publications directly on computer and they prefer to provide these products in multimedia environment (Gupta, 2000).

Digital Library
The idea of digital library is not new but recent developments have made the realization of this idea more of practical possibility than ever before. The availability of tools for easy access and production of the information has encouraged the traditional publishers to make their information available in electronic form bypassing the traditional publishing and bibliographic systems. This has led to the development of "Digital Library" (Upadhya, 2000).

With the advancing technology (Gupta, 2000) and the increasing cost of paper publications (Fox et al., 1995a) and shortage of floor space to house a library collection in the conventional form (Gupta, 2000), increasing use of computers, and decreasing budget, many libraries have to reduce their acquisition of books as well as their journals subscriptions (Fox et al., 1995a). The space problem compelled the user to discover the compact storage methods on different type of multimedia (Gupta, 2000). Similarly documents in electronic form can become more available and widely used because the cost of digital storage processing is going down (Fox et al., 1995a). This is the most compelling reason why a digital library – a library without wall is being created almost unconsciously.

History and Motivation
Futuristic views of libraries have long appeared in science fictions as well as in Bush's visionary statements (Bush, 1945) that foresaw and encouraged much work information retrieval and hypertext. Liclider (1965) laid out the clearest early call for digital libraries; he described a vast expansion of the publishing world as well as the need for knowledge organization, search, retrieval and many support activities. The early development of digital libraries resulted from three main factors:

  a. the human need for information,
  b. the growing volume of information accumulated in various communities, and
  c. advances in technology, e.g. networking, information storage and retrieval, hypermedia, and human computer interactions (Fox and Lunin, 1993).

In the USA one of the early statements was S.626 (Electronic Library Act of 1993), calling for "a system of state based electronic libraries, supporting the use of publicly available libraries, databases, and networks providing robust and reliable support services for search and retrieval."
The term "digital Libraries" became widely accepted as the result of many activities in this area in the period 1991-1993 (Fox, 1993). Perhaps corresponding to the growing interest in digital technology, such as digital networks, digital video and audio, and electronic publishing (Encyclopedia of Computer Science, 2000).

Definition
The term "Digital Library" has a variety of potential meanings, ranging from a digitized collection of materials that one might find in a traditional library through to the collection of

all digital information along with the services that make that information useful to all possible users (Leiner, 1998).

The digital library is library where the information is made available in electronic form and access to it is provided through computer and other media like local area network or Internet, etc. (Upadhya, 2000).

A digital library is a collection of information that is stored and accessed electronically. Digital library encompass two complementary ideas :

a. Digital library is a set of electronic resources and associated technical capabilities for creating, searching, and using information. In this sense there are an extension and enhancement of information storage and retrieval system that manipulate digital data in any medium (text, images, sounds; static or dynamic images) and exist in distributed networks. The content of digital library includes data metadata that describe various aspects of the data (e.g. representation, creator, owner, reproduction rights), and metadata that consist of links or relationships to other data or metadata. Whether internal or external to the digital library.

b. Digital libraries are constructed – collected and organized – by a community of users and their functional capabilities support the information needs and uses of that community. They are a component of communities in which individuals and groups interact with each other, using data, information, and knowledge resources and systems. In this sense they are an extension, enhancement, and integration of a variety of information institutions as physical places where resources are selected, collected, organized, preserved, and accessed in support of a user community. These information institutions include among others, libraries, museums, archives, and schools, but digital libraries also extend and serve other community settings, including classrooms, offices, laboratories, homes and public spaces.

The First idea emphasizes the fact that digital libraries are computer—based system constructed for people to use and that they are extensions of information storage and retrieval systems. The second emphasizes the belief that digital libraries should be constructed in a way that accommodates the actual tasks and activities that people engage in when they create, seek, and use information resources; in this sense they are an extension of physical environments. Both assert that digital libraries are sets of informant resources collected and organized on behalf of a community (National Science Foundation 1996).

Characteristics of Digital Libraries
Digital libraries requires undoubtedly digital technologies. The DL is a heterogeneous library which contains hard copy to online systems. The full-fledged digital library is one in which all the information is available in the digital form. This may however be difficult to achieve. Until recently, it consisted of mostly electronic documents, which are of reference type.

Now-a-days most of the documents are primary in nature. Further, queries received and processed are again of reference type in nature. The digital records from both local and foreign database may be used respectively for several purposes. Thus a considerable saving in effort, time and resources is possible (Rathore, 2000).

The Objectives of Digital Libraries

a. support diverse items such as multimedia, music, performing arts, and visual art collections

b. provide a paperless (digital) library of items that can be used by remote users currently

c. make library staff member more accessible to borrower for reference assistance by providing remote meeting capabilities

d. provide a modular system which can grow in function with minimal change to previously certified hardware and software (Whitepaper, 2001)

e. strengthen communication and collaboration between and among the research, business, government, and educational communities (CAN-LINKED, 1995)

Requirement of Digital Libraries

Digital library requires well-tested and proven information technologies including the multimedia kit. Much of the work in digital libraries is achieved through e-mail service, by participating in usernet (s), by accessing the database or servers through networks, like Internet. Locally developed databases will contribute lot to develop digital libraries. In order words, the components of digital libraries are :

- Local library system, with adequate PCs having LAN, local databases in machine-readable form. CD-ROMs etc. Provision to provide e-mail service, access to servers, and to remote databases etc.
- Networks, including the network of networks
- A variety of systems functions to coordinate, mange the entry and to retrieve data.
- Well-trained manpower

Today, the Internet, available as a result of the R&D in computer and communications, with hypertext-based navigating tools like the World Wide Web giving access to multimedia information has opened tremendous opportunities for disseminating better information services. The library profession responded positively to the concept of the digital library. Innumerable databases of research papers, pre-prints and technical reports are made available on the Internet already via Web Server for direct access and downloading. The world, digital library is a generic term that is used to broadly describe these development (Rathore, 2000).

Aspect of Digital Libraries

DLs are considered among the most complex and advanced forms of information systems. They involve many areas, not only in computer and information science but also in arts and humanities. Table 1 lists some of the related fields that contribute the advance of DLs.

Table 1. Areas of study, attributes, contents, features, issues and roles (adapted from Fox et al., 1995a

Abstracting Accessibility Agent Annotation Archiving Billing, Charging Browsing Catalogue Classification Clustering Commercial services Content conversion Copyright coherence Courseware Database Diagram (e.g. CAD) Digital video	Digitization Discipline-Level Library Distributed processing Document analysis Document model Economic Study Education support Electronic publishing Electronic study Filtering Geographic information system Hypermedia Hypertext Image processing Index Information retrieval	Intellect property rights Interactive systems Knowledge base Knowbot Library science Mediator Multilingual collection Multimedia stream playback Multimedia systems Multimodal National Library Navigation Object – oriented OCR OODB Support Personalization	Preservation Privacy Publisher library Repository Scalability Search Sociological study Storage Standards Subscription Sustainability Training Support Suability Virtual library Visualization World Wide Web

Storage of Dls and Interoperability

Documents are the heart of DLs. Without documents there would be no DLs (Encyclopedia of Computer Science, 2000). The digital library basically stores information in electronic form and manipulate large collection of these materials effectively. The digital library can be among the most complex and advanced form of information system because the digital library involves digital documents production, database management, information linking through hypertext, information retrieval, information filtering, instructional model, intellectual property right management, multimedia information service, reference service, information distribution and SDI etc (Upadhya, 2000). Thus digital library has heterogeneous and dynamic information environment, which deals with distributed collections of information (Fox and Sornil, 1999). This demands interoperability as essential characteristics. The interoperability issue is considered as a basic theme for discussion (Upadhya, 2000).

Fox and Marchionine (1998), observe that since the digital library incorporates linking different media into a single system, it requires interoperability and multilingual support. This could be:

    a. Technical interoperability
    b. Information interoperability
    c. Social interoperability

Technical interoperability is concerned with the hardware, networks, data types, and application compatibilities and protocols. Information interoperability incorporates content scope, language, metadata, naming conventions, semantics and user interface. The social interoperability address personal and organizational rights and responsibilities.
Multilingual support in a digital library environment is essential since this could be considered as cultural interoperability. This is also essential to overcome the language barrier.

At present the following issues are discussed in relation to further development of digital library and making it more of pervasive technology.

a. The technological issue is how to search and display desired selections from and across the large collections.
b. The practical approach to digital library focus on access cost and digitizing technology and how to mass manipulate the information on the Net.

The digital library thus stores large amount of documents of different media and format. It also enables to integrate the new media types and document formats into existing Digital Library (Bruce and Hsinchum, 1996).

Information Retrieval
Information retrieval involves the representation, organization, and storage of information, and access to it (Salton and McGill, 1983). DLs can be regarded as extended IR systems with multiple media and federation (Fox and Sornil, 1999). DLs may also be considered applications of IR. However, regardless of perspective, IR is essential for DLs to achieve high levels of effectiveness while affording ease of use to diverse communities. There is considerable research under way in IR related to DLs, dealing with multilingual processing, multimodal and structured documents, hypertext and hypermedia, searching of multimedia content, information visualization, handling distributed collections of complex documents, federated search standards, and information systems architectures (Encyclopedia of Computer Science, 2000).

Research in "nonstandard" types of documents, e.g. multimedia, multilingual, and hypertext, involves special kinds of languages, texts, complex structures, and indexing and searching techniques. However, the common requirement of all the types is content-based information access. The importance of multilingual information retrieval has been increased with the expansion of DLs (Encyclopedia of Computer Science, 2000).

Usability
"The key to the success of DL project is having usable systems" (Fox and Sornil, 1999). To achieve this goal, research in human – computer interaction focuses on modeling people's activities and needs in using DLs (Encyclopedia of Computer Science, 2000).

Property Rights and Access Control
The advancement of DLs depends on their successes not only in technology but also in legal, social, and economic contexts. We first discuss intellectual property issues of DLs and then

DLs in economic and social contexts. Copyrights was created to encourage authors to publish useful works by providing them suitable compensation for their creativity and labor. DLs introduce difficult and complex intellectual property issues because, with digital documents, multiple copies of a document can be viewed simultaneously since they can be easily transferred and duplicated. The collections may consist of priceless objects and valuable services; some methods to enforce property rights have to be applied to achieve a balance between author and publishers (interested in compensation) and user and library (interested in access to information). Some advocate subscription while others favor charges based on counts of use, unit-of-time-online, or bytes-delivered (Samuelson, 1995). The success of DLs, especially those supporting key applications like education, depends upon achieving a sensible balance among quality of service, usage, authority, and finances (Encyclopedia of Computer Science, 2000).

Economic and Social Implications

DLs introduce new relationships and impacts our world. They connect people together, supporting understanding of different cultures, heritages, and perspectives. New kinds of information are accessible to users. New possibilities of doing business result both directly from access to collections and indirectly from the content of the collections, e.g. advertisements. In publishing, DLs dramatically change the cost of production and dissemination, and libraries and publishers have to redefine their functions and purposes to survive in this environment. Publishers many adapt by performing functions such a selection of quality publications, editing, refinements, and promotions; they also may have some legacy information that DLs need to have in their collections. Libraries may preserve and provide access to historical records of works in their collection, may become the publishers of works prepared by their host institutions (e.g. when managing reports, as at http://www.ncstral.org, or theses, as at http://www. Theses.org), and may provide cross publisher catalogs (Samuelson, 1995). DLs blur the boundary between authors and readers; for example there are authoring and annotation tools that can be used for both publishing and reading. Copyright law, user privacy, and policies of collection development, management, preservation, and archiving need to be designed for effective use of DLs (Encyclopedia of Computer Science, 2000).

International Efforts

Thousands of digital libraries are emerging around the world, crossing all disciplines and media. Their sizes range from the small, serving local organizations, to the large, such as national libraries sharing highlights of history, cultural treasures, and accomplishments with the rest of the world (Fox and Sornil, 1999). Many countries have crated their own projects such as the Digital Libraries Initiative in the USA; the ERCIM program for the European community enhanced by projects in countries such as UK, France, and Germany; and related initiatives in Singapore and Japan. The international Digital Libraries Collaborative Research program of NSF will fund joint projects with these and other countries (Encyclopedia of Computer Science, 2000).

The world has diverse intellectual and cultural collection; connecting them together can lead to deeper understanding and wider cooperation. The major barriers are interoperability and multilingual collection processing. Some solutions are emerging, for example using Z39.50,

IP, and HTTP protocols plus international standards like SGML, SML (see MARKUP LANGUAGES), JPEG AND MPEG (see IMAGE COMPRESSION) (Fox and Marchionini, 1998).

Conclusion

In this era of the Internet and distributed, multimedia computing, new and emerging classes of information systems applications have swept into the lives of office workers and everyday people. New applications ranging from digital libraries, multimedia systems, geographic information systems, and collaborative computing to electronic commerce, virtual reality, and electronic video arts and games have created tremendous opportunities for information and computer science researchers and practitioners.

Digital libraries can meet the needs of user communities through a variety of services connected with complex collections and various structuring mechanisms for managing data and descriptions of that data. They involve not only computer and information science but also humanities and arts. Research and development in this field is being conducted actively in many communities. Each year workshops and conferences aim at establishing an agenda for investigation and creating new techniques and a consensus as to how to handle difficult problems. Important aspects that need to be carefully handled in DLs are document representations; collection creation, management, preservation, and archiving; usability; interoperability; intellectual property; social and economic implications; scalability; and supporting network infrastructure. Every issue needs to be handled in both local as well as global contexts, helping lead toward the concept of a world-wide digital library.

References

Bruce, S and Hsinchum Chen (1996). Building large scale digital library. Computer, 29(5). p 22-26.

Bush, V. (1945). 'As We May Think', The Atlantic Monthly, 176, 1. 01-108.

Drabenstott, Karen M (1994). Analytical Review of the Library of the Future, Washington, DC : Council Library Resources.
Encyclopedia of Computer Science (2000). Ralston, Anthony, Reilly, Edwin D. and Hemmendinger, David (eds). 4th edn. London : Nature Publishing Group. pp 576-581.

Fox, E.A. (1993) Digital Library Source Book. http://fox.cs.vt.edu/DLSB.html

Fox, E.A., and Lunin, L.F. (1993). Introduction and Overview to Perspectives on Digital Libraries. Journal of American Society for Information Science, 44, 8, 480-491

Fox, E.A., and Sornil, O (1999). Modern Information Retrieval, Harlow, UK : Addison Wesley Longman.

Fox, Edward A and Marchionini (1998). Towards a worldwide digital library. Communication of the ACM, 41(4). pp. 29-32

Gupta, R.C. (2000) Digital Libraries : An Emerging Reality. In : Babu, T. Ashok, Ramaiah, L.S.; Saxena, S.C. and Bedi, D.S. (eds). *Vision of Future Library and Information Systems*. New Delhi : Viva Books Pvt. Ltd. pp. 225-229

Leiner, Barry (1998). The Scope of the Digital Library. Paper presented in the Dlib Working Group on Digital Library Metrics, at Stanford University, USA: January 7-8. Revised October 15, 1998. Also available in http : // www.dlib.org/metrics/paper/dig-lib-scope.html

National Science Foundation (1996). Social Aspects of Digital Libraries. *Final Report*. Report presented in UCLA-NSF Workshop on Social Aspects of Digital Libraries, University of California, Los Angeles, USA: February 15-17.

Rathore, R.S. (2000) Digital Libraries : A Boon for Resource Sharing and Networking. In : Babu, T. Ashok, Ramaiah, L.S.; Saxena, S.C. and Bedi, D.S. (eds). *Vision of Future Library and Information Systems*. New Delhi : Viva Books Pvt. Ltd. pp. 225-229.

Salton, G., and McGill, M. J (1983). Introduction to Modern Information Retrieval. New York : McGraw-Hill.

Samuelson, P. Copyright and Digital Libraries, Communications of the ACM, 38, 3, 15-21 Upadhya, Padma V. (2000) Digital Library Approach to Geographic Information System (GIS). In : Babu, T. Ashok, Ramaiah, L.S.; Saxena, S.C. and Bedi, D.S. (eds). *Vision of Future Library and Information Systems*. New Delhi : Viva Books Pvt. Ltd. pp. 225-229.

Whitepaper (2001) System Level Specification for the Digital Library System. The University of Texas at Dallas. July. Ver. 1.1