DESIDOC Bulletin of Information Technology, Vol. 16, No. 1, January 1996, pp. 31-39

A Course in Digital Libraries

A Neil Yerkey* and Corinne Jorgensen
State University of New York at Buffalo
Buffalo, NY 14260, USA

ABSTRACT

The paper describe the objectives, topical content, and reasons for developing a course on digital libraries. The Internet is changing the way librarians view information sources. The idea of the library as a place is being questioned while the "library without walls" is rapidly emerging. The course provides a broad overview of technical and professional issues related to digital libraries. It leads students through a theoretical discussion of various facets, and gives practical experiences with tools necessary for creation and use of information resources in networked digital formats.

1. THE ACADEMIC SETTING

The State University of New York at Buffalo (SUNY at Buffalo), of which the School of Information and Library Studies (SILS) is a part, is ranked among the top public research universities in the US. It has more than 27,000 students and 4500 faculty and offers 91 undergraduate, 99 master's and 90 doctoral degree programs. SUNY at Buffalo is a member of the prestigious Association of American Universities.

The School of Information and Library Studies (SILS) is one of about 15 major academic units within SUNY at Buffalo. Its primary function is to prepare graduates for positions in all types of libraries and information agencies. In addition, it grants a Post-Master's Certificate and has a joint doctoral program with the School of Education. It is moving toward establishing its own doctoral program in library and information studies, and has recently established the Centre for Applied Research in Library and Information Science to co-ordinate and conduct research of practical benefit to libraries. It currently has about 252 students in its master's level program. SILS is fully accredited by the American Library Association.

* *Communication should be addressed to :*
A Neil Yerkey, School of Information and Library Studies
379 Baldy Hall, SUNY at Buffalo, Buffalo, NY 14260, USA

One of the foundational philosophies of SILS has been to develop a curriculum that reflects the merging of the science and technology of information, print, non-print, and computer media within librarianship. It was one of the first library schools to offer courses on online searching when that technology was in its infancy. Besides teaching students how to use technology within libraries, SILS also recognised the need to train people to understand, work with, and contribute expertise to the 'informatization' of society outside the library walls.

2. CHANGES IN LIBRARIANSHIP

Librarians have been moving away from the idea of the library as a storehouse of information, and toward the idea of libraries as places where information is managed [1]. Librarians embraced 'newer media' such as films, filmstrips, video and audio cassettes. They realised that information was contained in many forms other than print on paper. Even so, the profession until recently viewed the library as a place where information 'containers' are acquired, organised, shelved, and retrieved.

More recently, librarians began to use online search database systems, such as Dialogue and Medline, to extend the boundaries of the library's resources. They embraced remote, shared systems like OCLC, and developed online public access catalogues (OPACs) with connections to the OPACs of other libraries. The walls of the library began to be less solid. Still, until recently, librarians were mostly concerned with containers. They thought of information as a thing and efforts were directed toward organising the things that contain information. Even the goal of newer techniques like online searching was ultimately to help retrieve things like articles, reports, and books [2]. Only very recently, due to the emergence of global networks, have librarians had the capability to conduct real information retrieval, as opposed to container retrieval.

3. THE INTERNET

The Internet is changing the way we view information sources. Not only is the idea of the library as a place being questioned, but librarians are finding they must learn how to organise, manage, use, and provide access to containerless data and ideas. Personal computers are now usually part of a network of computers and often part of a network of networks, spanning the globe. Until recently, the few who were experimenting with networks were computer scientists, government-sponsored researchers, and hobbyists who could brave the complexities of the newly emerging global village. For most it was out of reach, economically and practically.

Now, interconnectivity is much more widespread. Computers are less expensive, people are becoming more computer literate, and systems are becoming simpler and more accessible. These changes have far-reaching implications for the workplace and society. Computers are changing the way we communicate and how we think about intellectual property and the ownership of information. It is forcing us to re-examine our ideas about information, libraries, and librarianship.

The Internet is not the first or only computer network. It is certainly not the best organised, or the easiest to use. It can best be described as a sprawling mass of computer connections that

somehow works. It is the first attempt at a "Global Network," connecting millions of computers in more than 50 countries (100 more countries have some kind of indirect connection). It is actually not a single network, but many regional networks connected by a backbone of several supercomputers which communicate via satellite and fiber. More importantly, the Internet is a community of 20 million people, and a vast library of information. People around the world are able to contribute their thoughts and knowledge to the community, and everyone else can draw off their knowledge [3].

Some of the services provided on the Internet are:

- **E-mail :** to communicate with others by sending private messages to a specific person at a specific computer address.
- **Telnet :** To connect and log onto a remote computer, providing access to online public access catalogues, databases, information sources of various kinds, and software on other computers.
- **Ftp :** File Transfer Protocol, used to send and get files of text, computer programs, pictures, games and music to and from remote computers.
- **Mailing lists :** Mailing lists which allow people to exchange e-mail on topics of interest.
- **Electronic documents :** Articles, books, news items, and the like available as "full text" documents.
- **OPACs :** Connections to online library catalogues around the world.
- **Bulletin boards and newsgroups :** Open or moderated discussions on specific topics.
- **Archives :** Holdings of bibliographies, documents, newsletters, software, and the like.
- **Databases, online search services :** Indexed literature and other information resources with search capabilities.
- **Freenets :** Community information networks.

The Information Age, talked about for many years, is now moving upon us with increasing speed, fuelled by the merging of computers and telecommunication technology.

Some effects of this revolution on libraries are :

- Libraries without walls—"virtual libraries."
- Availability of information resources to end users with less need for professional intermediaries.
- Availability of electronic journals, conferences, books, discussions, and newsgroups.
- Less reliance on collection development and maintenance and more on resource sharing.
- More reliance on having information and data retrieved and packaged on demand.
- Increasing attention being paid to "containerless" data and ideas.
- Global connectivity—the Internet.

4. DIGITAL LIBRARIES

Recently, the idea of the "digital library" has moved to the forefront of discussion and research. This was due in part to the activities of the Library of Congress in helping to establish a National Digital Library Federation. In May 1995, the Commission on Preservation and Access

(a private organisation whose mission is to develop and support collaboration among libraries to ensure access to and preservation of resources in all formats), the Library of Congress, and 14 other research libraries signed the National Digital Library Federation Agreement [4]. The purpose was to "bring together–from across the nation and beyond–digitised materials that will be made accessible to students, scholars and citizens everywhere" (Ref 4, p. 251). The participants agreed to establish a collaborative management structure, develop a co-ordinated approach to fund-raising, and formulate cooperative selection guidelines.

The Library of Congress has long maintained many information resources on the Internet. These include electronic exhibits, the American Memory Project, ftp sites, country studies, LC Online, indexes to Federal Government resources, and a link to the Library's Thomas Legislative System. More recently it has embarked on a plan to digitise its collections relating to American history. It has also been addressing copyright issues through development of an electronic copyright management program [5].

Other libraries have embarked on similar projects. The British Library began an "Initiatives for Access" project in 1993 to investigate hardware and software platforms for the digitization and subsequent networking of a range of library materials. It includes a patent information system, "Electronic Beowulf" digital manuscript, photographic collections, and an OPAC [6].

On September 27, 1994, the National Science Foundation (NSF), the Department of Defence Advanced Research Projects Agency (ARPA), and the National Aeronautics and Space Administration (NASA), through a joint initiative, announced the funding of several digital library development projects, mostly among academic institutions. The focus of the digital library initiative is to dramatically advance the means to collect, store, and organise information in digital forms, and make it available for searching, retrieval, and processing via communication networks. Each project is described briefly below. (More information may be found at the World Wide Web addresses (URL's) shown.) The differences in how the grantees define "digital library" are especially instructive.

- **The University of Michigan** project defines digital libraries as "electronic libraries in which large numbers of geographically distributed users can access the contents of large and diverse repositories of electronic objects. Electronic objects include networked text, images, maps, sounds, videos, catalogues of merchandise, and scientific, business, and government datasets. It will concentrate on earth and space sciences (URL:http://http2. sils.umich.edu/UMDL/HomePage.html)."

- **Stanford University** is developing the enabling technologies for a single, integrated, virtual library, allowing users to access information through interfaces that hide the unimportant details of diversity of materials and provide ways to navigate and manage the information space in a consistent and unified way. Stanford's project will include a mix of virtual and traditional libraries (URL: http://www-diglib.stanford.edu/diglib).

- **The University of California** at Berkeley is producing a prototype digital library with a focus on environmental information. Its goal is to develop a system in which large

numbers of geographically distributed users can conveniently access the entire contents of very large and diverse repositories of electronic objects. These repositories will exist in locations physically near or remote from the users, and will contain objects comprising text, images, maps, sounds, full-motion videos, merchandise catalogues, scientific and business data sets, as well as hypertextual compositions of such elements (URL: http://elib.cs.berkely.edu).

- **The University of Illinois** believes the international network will evolve from distributed computer nodes to distributed information sources supporting object interaction. Users will browse the Net by searching digital libraries and navigating relationship links, as well as share new information within the Net by composing and publishing new objects and links. The Net will thus appear as interconnected spaces of information objects (URL:http://www.grainger.uiuc.edu/dli).

- **Carnegie Mellon University** is developing an online digital video library, allowing users to access, explore and retrieve science and mathematics materials from video libraries. This project is more concerned with connecting to an existing storehouse and accessing the information in television videos (URL:http://fuzine.mt.cs. cmu.edu/im/iformedia.html).

- **The University of California at Santa Barbara** is working on a project called Project Alexandria which is primarily concerned with maps and geographical images and geographical information systems (URL:http://alexandria.sdc.ucsb.edu).

These project descriptions demonstrate the emergence of a new paradigm for information storage, access, and delivery– that of the "virtual" or "digital" library. As can be seen, the terminology is by no means settled. Drabenstott [7] points out that the literature contains words like "digital library," "electronic library," "virtual library," "library without walls," and "bionic library " to refer to the library of the future. Additionally, these terms have different meanings among the different interested parties involved:

> To some it suggests computerisation of traditional libraries. To others, it calls for carrying out of the functions of libraries in a new way, encompassing new types of information resources; new approaches to acquisition, cataloguing, interaction, etc. To computer professionals it is a distributed text-based or multimedia information system. Some see it as linked multimedia objects that encompass humanity's store of information. Some see digital libraries as the space in which people communicate, share, and produce new knowledge and knowledge products [8].

This statement reflects two different but complimentary ways of thinking about the concept. One is to think in terms of digital products in traditional libraries. That is, the definition of a digital library is a more or less traditional library that uses CD ROMs, online searching, online public access catalogues, e-mail, and other electronic products and tools to serve its clientele. Indeed, libraries are increasingly using digital products, tools, and resources to serve very traditional functions. Part of the ever-growing body of literature is concerned with the issues revolving around this side of the concept [9].

The other approach is to think more in terms of virtual libraries. This approach focuses not on traditional libraries in library buildings, but rather on online

information resources used by people in all kinds of settings. Virtual libraries sometimes "look" like traditional libraries. An example is the Internet Public Library developed by the University of Michigan School of Information and Library Science (URL: http://ipl.umich.edu). More typically, virtual libraries are personal, ad hoc creations, developed to meet the informational and recreational needs of one individual or a small group. They might be related to a specific project or developed and maintained throughout a lifetime of research and interest by an individual. Or, virtual libraries might be developed for public use in much the same way public libraries have been developed.

There are, however, commonalties among the variety of projects and the multiple definitions being offered for the "digital" or "virtual" library:

- Geographically distributed users.
- Large and diverse collections.
- Information represented by a variety of electronic objects.
- "Seamless" access.

McClure has pointed out that "the virtual library would allow users to seamlessly connect and interact with information with no regard to geographic location or time" [10]. No buildings here, no containers, no staff. Librarians of the digital library of the future should assume the role of information manager, both in creating and facilitating access to a variety of information objects across multiple platforms.

Libraries as places will continue to serve useful informational and communal functions. Also, containers of information will be around for a long time, and they will continue to be selected, acquired, catalogued, indexed, and referred to. In fact, among the definitions and descriptions of digital libraries, Nunberg describes the electronic library as :

> An aggregation of catalogues, lists, and indexes of documents of every imaginable type, organised according to myriad schemes of classification, and linked and cross-indexed for searching, so that they come to behave as a single database in which the lines between individual collection and catalogues are blurred [11].

5. COURSE DESCRIPTION

Because we believe traditional libraries will be part of the mix, the course in Digital Libraries developed at SILS includes some discussion of the impact of digital products in traditional library settings but emphasises the changes being instituted by the creation of virtual libraries on existing cultural, social, organisational, economic, intellectual, and technical infrastructures for information delivery. Students taking the course will acquire the skills necessary to develop, organise, use, and evaluate the electronic information products which comprise digital libraries.

As a working definition for the course, we define digital libraries as electronic libraries in which large numbers of geographically distributed users can access the contents of large and diverse repositories of electronic objects. Electronic objects include networked text, images, maps, sounds, videos, catalogues of merchandise, and scientific, business, and government datasets. They also

include hypertext, hypermedia, and multimedia compositions.

The course provides a broad overview of technical and professional issues related to digital libraries. It provides both theoretical discussion of various facets, and practical experiences with the tools necessary for creation and use of information resources in networked digital formats.

This course was offered as a three credit-hour, 15 week course at the Master's level. Students were expected to already have an understanding of basic Internet resources such as e-mail and bulletin boards or listservs. Most of the class time was spent in lecture, discussion, and presentations by outside experts. Some class time was used for hands-on instruction and demonstrations, and there were a few field trips to view digital projects in the area.

The main project for the course consisted of working in small groups to develop a digital product or virtual information structure. The projects consisted of creating World Wide Web pages for area libraries or other information agencies, developing information areas on the Buffalo Free-Net, creating digital multimedia programs, and the like. Students also wrote individual papers describing the process, and each group conducted a class demonstration and presentation of its projects.

Activities and assignments were designed to help students:

- Identify and understand professional issues related to digital libraries, such as the place of library and information centres in a digital environment, ownership of information, censorship and fair use, collection development, and information mediation.
- Understand technical aspects related to storage, transmission, and retrieval of digital information, such as networking, protocols, client/server environments, digital storage formats, compression, retrieval, and presentation.
- Develop necessary skills and knowledge of tools for creation and effective use of networked, multimedia applications.
- Examine and develop library and public policy concerning collection development, security and privacy, partnerships and collaboration, information have-nots, intellectual freedom, and other social and cultural issues in a networked environment.

6. COURSE TOPICS

The main topics for the course are outlined below. Technical issues and vocabulary are defined in the course, but students desiring a more in-depth treatment of such material are encouraged to take appropriate courses within other schools and departments of the university.

Topic 1: Digital Libraries

An overview of the concept of digital libraries and the changes taking place in the information age : Vision and overview, definitions, National Science Foundation initiatives and responses, implications for librarians, and demonstrations of digital projects.

Topic 2: The Internet

History of the Internet and evolution of its resources : History of networks, types of resources and tools available,

and trends in the development of information browsers and organisers.

Topic 3: Basic Internet Tools

Hands-on experience with Internet tools and resources : Telnet, ftp, gopher, free-nets, browsers, and other virtual structures.

Topic 4: Network Technology

Non-engineering discussion of local and wide area networks : Networking systems, types of wiring and configurations, protocols, security, packet switching, interoperability and standards.

Topic 5: Client/server Technology

Discussion of trends in moving from dumb terminal processing to a client/server environment : Standards, interfaces, capacity, costs, administrative issues, access and security.

Topic 6: Document Handling

Discussion and demonstrations of how to create, store, and transmit digital products : Document generation, storage and transmission formats, data capture, scanning techniques, compression, conversion, and preserving digital objects.

Topic 7: Hypermedia and Multimedia

Discussion of hypertext and hypermedia techniques : Incorporation of graphics, audio, hypertext, Geographic Information Systems [GIS], and multimedia information into digital products, and architectures, reference models and standards.

Topic 8: Research and Scholarship Issues

Discussion of how digital libraries are changing research and scholarship : Issues of validity and reliability of electronic information, problems of dating, updating and timeliness of information.

Topic 9: Resource Development & Analysis

Discussion of collection development within a digital library setting : How to analyse system and resource usage, how to evaluate effectiveness, how to plan for a digital environment, and issues of cooperative collection development by libraries.

Topic 10: Ownership

Issues concerned with copyright and economies in a digital world: Intellectual property rights, authoring and electronic publishing, library and publisher's economic issues, costs and cost recovery of digital products.

Topic 11: Social/cultural Issues

Discussion of censorship of networked information and information have-nots : Censorship and intellectual freedom, problems of disparity of information access, role of community (free) information systems.

Topic 12: Indexing, Search, and Retrieval

Discussion and demonstration of how to find and organise networked information : Knowledge and information representation, bibliographic organisation and control, indexing, surrogation, information filtering, retrieval, browsing and information exploration and automatic techniques.

Topic 13: User Interfaces

Discussion and demonstration of how to best present information to end users : Interface/display issues, user needs in data visualisation and representation, evaluation methods and user testing.

Topic 14: Library of the Future

The role of libraries and librarians in the information age : Librarian mediation in a

virtual environment, issues of a paperless society, and library support issues.

Topic 15: Partnerships and Collaboration

Discussion of new opportunities for information creation and dissemination : Collaborative environments, networked information discovery, education, learning and related applications, business, medical and reference uses.

The course outline presented above attempts to balance breadth of coverage of the many issues involved in creating digital libraries with exploration of particular issues and techniques through additional readings, assignments and projects. The technology continues to change rapidly; therefore the course emphasises acquainting students with the broader organisational, social, economic, and cultural contexts within which decisions concerning digital information products must be made, rather than focusing on specific technical implementations.

Peter Graham reminds us that libraries, digital or traditional, must be committed to organising, storing and providing information for periods of time longer than human lives. They are not simply networks full of databases nor buildings full of books; they should be organisations designed for the long term [12]. We believe this perspective is necessary if the digital libraries of the future are to be designed to truly serve the needs of information seekers rather than to simply seek the limits of what is technically feasible.

7. REFERENCES

1. Braude, R.M.; Florance, V.; Fisse, M.; & Fuller, S. The organisation of the digital library. *Academic Medicine,* 1995, 70(4), 286-291.

2. For a discussion of this issue see Buckland. Information as thing. *J. Am. Soc. Info. Science,* 1991, **42**(10), 351-360; Yerkey, N. Information as thing, (ltr) and Buckland, rejoinder. *J. Am. Soc. Info. Science,* 1991, **42**(10), 758.

3. Krol E. & Hoffman E. FYI: What is the Internet? User Services Working Group of the Internet Engineering Task Force, May, 1993 (Online document: ftp://ftp.univie.ac.at/netinfo/rfc/rfc1462 txt).

4. Lamolinara, G. A. New federation: LC and others sign National Digital Library Federation Agreement. *Library of Congress Information Bulletin,* 1995, **54**(12), 251.

5. Becker, H.S. Library of Congress digital library effort. *Communication of the ACM,* 1995, **38**(4), 66.

6. Purday, J. The British Library's initiatives for access projects. *Communication of the ACM,* 1995, **38**(4), 65.

7. Drabenstott, K. M. Analytical review of the library of the future. Council on Library Resources, Washington, D.C., 1994.

8. Fox, E.A.; Akscyn, R.M.; Furuta, R.K. & Leggett, J.J. Digital libraries. *Communication of the ACM,* 1995, **38**(4), 23-28.

9. Saunders, L.M. (Ed). The virtual library: visions and realities. Meckler, New York, 1993.

10. McClure, C.R., McKenna, M., Moen, W.E., & Ryan, J. Toward a virtual library: Internet and the National Research and Education Network. *In* The Bowker annual of library and book trade

information. Reed Publishing, New York, 1993. pp. 25-45.

11. Nunberg, G. The place of books in the age of electronic reproduction. *Representations,* 1993, **42** (Spring), 13-37.

12. Graham, P. The digital research library: Tasks and commitments. *In* Digital Libraries '95; The second annual conference on the theory and practice of digital libraries, 11-13 June 1995, Austin, Texas, 1995. p. 95.