

DIGITAL LIBRARIES: IMPACT ON SCIENCE, MEDICINE, AND LIBRARIES

**Ruth Hodges
South Carolina State University**

ABSTRACT

Congress's Electronic Library Act of 1993, and the President and Vice-President in their report on "Technology for America's Economic Growth", called for the development of digital libraries. Workshops by the National Science Foundation (NSF) and others led to defining a research agenda for digital libraries, which was identified as a national challenge in the Information Infrastructure Technology Application component of the U.S. High Performance Computing and Communications Program. In late 1993 the National Aeronautics and Space Administration (NASA), NSF, and the Advanced Research Projects Administration (ARPA) jointly sponsored a call for proposals on the topic of "digital libraries." Focus of the project was to dramatically advance the means to collect, store, and organize information in digital forms, and to make it available for searching, retrieving, and processing via communications networks in a user-friendly way. Six projects, centered at Carnegie-Mellon University, the University of California-Berkeley, University of California-Santa Barbara, Stanford University, University of Illinois Urbana-Champaign, and the University of Michigan, were funded under this initiative. Project multidisciplinary collaboration occurred between librarians, engineers, computer scientists, and others.¹ This poster focuses on current developments in digital library research and technology; products and applications of digital library technology in science, medicine, and libraries; issues resulting from the development of digital libraries such as copyright, Internet quality control, etc., and to what extent are Historically Black Colleges and Universities (HBCUs) involved in digital library technology.

WHAT IS A DIGITAL LIBRARY?

There are many definitions of a "digital library." Terms such as "electronic library" and "virtual library" are often used synonymously. The elements that have been identified as common to these definitions are:⁷

- The digital library is not a single entity
- The digital library requires technology to link the resources of many
- The linkages between the many digital libraries and information services are transparent to the end users
- Universal access to digital libraries and information services is a goal
- Digital library collections are not limited to document surrogates: they extend to digital artifacts that cannot be represented or distributed in printed formats.

In 1995 this broad definition was adopted by the Association of Research Libraries, a major library organization mostly comprised of large academic libraries.

The components of a digital library include assemblage of electronic data, cataloging and indexing mechanisms, tools for locating, searching, and browsing data collections, mechanisms to retrieve and potentially process data from remote and distributed locations, and interface tools that allow tasks to be performed easily by non-expert users. Because each of these components requires expertise from several disciplines, collaborations are the norm—as evidence by several existing projects.⁸

Individuals proficient in: Cognitive analysis, psychology, human computer interaction techniques, graphic arts, specification languages, algebra, user interface programming and management.¹

IMPACT ON SCIENCE AND MEDICINE

In 1993 the NASA, NSF, and ARPA jointly sponsored a call for proposals on the topic of “digital libraries.” The focus of the project was to dramatically advance the means to collect, store, and organize information in digital forms, and to make it available for searching, retrieving, and processing via communication networks in a user-friendly way; six projects were funded under this initiative (see table 1). The number of accessible collections has continued to grow with the addition of global change datasets, climate and weather databases, human genome data, and the social sciences (see table 2).¹ Because of the complexity of these projects, they each include multidisciplinary teams consisting of engineers, librarians, statisticians, computer scientists, and others. As documents are integrated into very large collections covering an entire scientific domain, links among the documents become increasingly important to help with searching and browsing. Who more than librarians have the expertise and experience to identify the important linkages between documents that should be incorporated within catalog entries.

IMPACT ON LIBRARIES

Technological advances have resulted in changes in libraries. Digital technology empowers various libraries, serving various populations, to link together and even merge. Moreover, many libraries have created digital libraries to further expand access, increase usability and effectiveness, and establish entirely new ways for individuals to interact with information. Some envision a “hybrid” library—one that combines traditional print publications and new digitized information. Others view digital technology as a means of preserving deteriorating manuscript and rare books and have joined the “digital bandwagon”. For example, Cornell University has put more than 2,000 19th and early 20th century books in electronic format; conservators of the Library of Congress, the British Library in London, and the Vatican Library in Rome all are digitizing their most treasured volumes. The Library of Congress expects to have 5 million books and images digitized by the year 2000.¹¹

HISTORICALLY BLACK COLLEGES AND UNIVERSITIES

Although technological infrastructure may be considered modest for many Historically Black Colleges and Universities (HBCUs), most of them have significant archival collections. Howard University has its world renowned Moorland-Spingarn collection which attracts scholars nationwide to research. These collections contain archival materials on Black people throughout the Diaspora. For example, the collection includes original papers, correspondence, etc., of Charles Drew (physician), Montague Cobb (anatomist), Frederick Douglass (abolitionist), Kwame Nkrumah (president and revolutionist), Alan Locke (writer and Rhode Scholar), E.E. Just (zoologist), and many others.

South Carolina State University’s Historical collection contains original documents of the Palmetto Medical, Dental and Pharmaceutical Association, a South Carolina African American organization originating in 1896. Claflin University, a neighboring college, houses minutes of the Methodist church dating back to the 1880s. The digitalization of archival documents at HBCUs would expand the use of these documents and promote resource sharing among institution—ideally similar to Georgia’s Galilee system.

As digital libraries, archival collections and other library resources would provide value added sources of information. This information could be disseminated and expanded nationwide to students and researchers studying science, medicine, history, genealogy, etc., thus promoting research and development nationwide.

ISSUES RELATED TO DIGITAL LIBRARIES^{3, 15, 4}

1. Copyright — how do we make materials widely available digitally while protecting the library rights of artists, writers, and publishers?
2. Interoperability — a popular vision of the digital library is that a library user at one location using one interface could seamlessly search the digital collections of hundreds of libraries. This requires digital libraries nationwide and internationally to comply with certain key standards. **The key technological issues are how to search and display desired selections from and across large collections?**
3. Selection — what should we digitize?
4. Economics — where will we get the money to maintain current projects and fund future ones?
5. Will what is available today be available tomorrow in this rapidly changing technology?

CONCLUSION

Given the infinite possibilities of digital technology, research and development in this area proceed to new heights. HBCUs like other institutions need to strengthen their technological infrastructure to ensure that digital libraries are fully functional—“no need to have the technology if it doesn’t work.” Institutions need to develop digital libraries with universal search engines, with user friendly systems, and within the regulations of copyright. Accordingly, institutions need to increase their library’s budgets to enable them to subscribe to more value added Internet resources, e.g., online journals, indexes, reference books, etc.—by doing so institutions expand access to more quality Internet resources and other research materials for students, faculty, and staff—thus reducing or eliminating geographic and temporal barriers.

Table 1: Six Digital Libraries Projects^{4,11,15}

Project Name	Project Description
<p>Carnegie-Mellon University Project Title: <i>Informedia Digital Video Library</i></p>	<p>Investigating how multimedia digital libraries can be established and used (video, audio, images and text); research is in the area of speech recognition, image understanding, and natural language processing supports. . .knowledge base and search retrieval.</p>
<p>University of California-Berkeley Project Title: <i>Environmental Planning and Geographic Information Systems</i></p>	<p>Goal is to develop the technologies for intelligent access to massive, distributed collections of photographs, satellite images, maps, full text documents, and multivalent" documents; research includes computer vision, databases and information retrieval, user needs assessments, etc.</p>
<p>University of California-Santa Barbara Project Title: <i>The Alexandria Project: Spatially-Referenced Map Information</i></p>	<p>Explores problems related to a distributed digital library for geographically referenced information; comprised of digital maps, text, space shuttle and satellite images, and historic aerial photography of California.</p>
<p>Stanford University Project Title: <i>Interoperation Mechanisms among Heterogeneous Services</i></p>	<p>Focuses on interoperability, i.e., how disparate databases can be treated as one by the user (Tennant,1997). The network will include not only books and journals, but data from private collections and scientific studies. The system employs software enabling people to browse through disparate sources at the same time.</p>
<p>University of Illinois Urbana-Champaign Project Title: <i>Federating Repositories of Scientific Literature</i></p>	<p>Will make available full-text journals, magazines, and scientific literature primarily for engineers. Documents mostly in SGML (Standard Generalized Markup Language) and Adobe Acrobat. This system will provide people . . . at 10 large Midwestern universities remote access to full-text and pictures of tens of thousands of documents.</p>
<p>University of Michigan Project Title: <i>Intelligent Agents for Information Location</i></p>	<p>Sophisticated software programs are being developed which will allow librarians and users to search more efficiently for information on the subjects of earth and space science; JSTOR (Journal Storage Project), Humanities Text Initiative, etc.</p>

Table 2: Other Digital Libraries Projects^{5,9,10,13-16}

Project Name	Project Description
New York Public Library Project Title: Digital Library Collection	Includes images and texts files from the Schomburg Center for Research in Black Culture pertaining to women, dance, theater, music, etc.; and the NYPL finding aids.
Indiana Music Library Project Title: VARIATIONS Project	digitalization of audio and video materials related to music.
University System of Georgia Project Title: Galilee	Statewide library resource sharing among Georgia's private and state universities, public libraries, technical schools, and public schools.
Library of Congress Digital Collections Project Titles: American Memory; and Thomas	Historical Collections, and Legislative Information on the Internet, respectively.
NASA Project Title: Digital Library Technology Project (DLT)	Provides support for those developing "technologies to facilitate public access to NASA data via computer networks.
National Radio Astronomy Observatory Library (NRAO)	Database of all preprints received at NRAO from 1986 forward, with citations to published papers
NCSA Astronomy Digital Library	Collects astronomical, research-ready images and make them available to the astronomical community and general public.
Project Gutenberg	Electronic text files.
Perseus Project	Multimedia library of information about Archaic and Classical Greece: Literature, history, art, and archaeology.
National Library of Medicine Project Title: Visible Human Project; and GeneBank	Anatomically 3-D representation of the male and female human body; and NIH's genetic sequence database of all publicly available DNA sequences, respectively.
Virginia Tech Project Title: Networked Digital Library of Theses and dissertations.	An initiative consisting of 39 institutions to increase the availability of student research by digitizing masters theses and doctoral dissertations.
University of Iowa Project Title: The Virtual Hospital	A medical multimedia database containing materials for patient care, patient education, and continuing medical education.

REFERENCES

1. Beard, Kate. "Digital spatial libraries: a context for engineering and library collaboration." *Information Technology and Libraries*. 14:79-84, 1995.
2. Buildings, Books, and Bytes: Public Visions, Private Reflections. <http://www.benton.org/Library/Kellogg/chapter1.html#technolo> (7 July 1998).
3. Chepesiuk, Ron. "The Future is here: America's libraries go digital", *American Libraries* 28(1): 47-50, 1997.
4. Digital Libraries. **Error! Reference source not found..** (28 August 1998).
5. Digital Library Projects. <http://www.texshare.edu/Services/Professional/digital.html>. (26 August 1998).
6. Digital Projects of the UC Berkeley Library: Summary July 1998. <http://sunsite.berkeley.edu/R+D/ucblibrary.html> (28 August 1998)
7. Drabenstott, Karen M. Analytical review of the library of the future, Washington, DC: Council Library Resources, 1994.
8. Fox, E. A., D. Hix, L.T. Nowell, D.J. Brueni, W.C. Wake, and L. Heath and D. Rao. Users, user interfaces, and objects: Envision, a digital library", *Amer. Soc. Infor. Sci.* 44(8): 480-91, 1993.
9. Galileo: Georgia online learning online. <http://www.galileo.peachnet.edu/Homepage.cgi>. (4 September 1998).
10. Hart, Michael. What is Project Gutenberg: History and philosophy of the project. <http://promo.net/pg/history.html> (4 September 1998).
11. Lipkin, Richard. "The library that isn't there; digital libraries transform books, photos, and videos into bits and bytes", *Science News* 147(22):344-346, 1995.
12. National Library of Medicine. National Institutes of Health. National Center for Biotechnology. <http://ncbi.nlm.nih.gov/> Revised 25 Aug 98 (4 September 1998).
13. Networked Digital Library of Theses and Dissertations. <http://www.ndltd.org/> (4 September 1998).
14. New York Public Library. Digital Library Collections. <http://www.nypl.org/> (4 September 1998).
15. Tennant, Roy. "Digital Libraries: The grand challenges", *Library Journal* 122: 31-32, 1997.
16. University of Iowa. The Virtual Hospital. <http://indy.radiology.uiowa.edu/> (10 September 1998).