

WHAT ARE DIGITAL LIBRARIES? COMPETING VISIONS

Christine L. Borgman, Professor & Presidential Chair in Information Studies
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095
cborgman@ucla.edu

October 23, 1998

Final Revised Version for Information Processing & Management

Special Issue on Digital Libraries

Gary Marchionini & Edward A. Fox, Issue Editors

ABSTRACT	2
INTRODUCTION	3
PERSPECTIVES ON DIGITAL LIBRARIES	4
Research vs. Practice	5
Community Building Efforts	7
Framing the Issues	8
DEFINING DIGITAL LIBRARIES	11
Digital Libraries as Content, Collections, and Communities	11
Research-oriented Definitions	12
Defining Elements of Digital Libraries	16
Digital Libraries as Institutions or Services	17
Digital Libraries as Databases	19
INTO THE DIGITAL FUTURE	20
SUMMARY AND CONCLUSIONS	22
ACKNOWLEDGEMENTS	24
NOTES	25
REFERENCES	26

ABSTRACT

Research and practice in digital libraries (DL) has exploded worldwide in the 1990s. Substantial research funding has become available, libraries are actively involved in DL projects, and conferences, journals, and online news lists proliferate. This article explores reasons for these developments and the influence of key players, while speculating on future directions. We find that the term "digital library" is used in two distinct senses. **Many** **R**esearchers view digital libraries as content collected on behalf of user communities, while practicing librarians **typically** view digital libraries as institutions or services. Tensions exist between these communities over the scope and concept of the term "library." Research-oriented definitions serve to build a community of researchers and to focus attention on problems to be addressed; these definitions have expanded considerably in scope throughout the 1990s. Library community definitions are more recent, and serve to focus attention on practical challenges to be addressed in the transformation of research libraries and universities. Future trends point toward the need for extensive research in digital libraries and for the transformation of libraries as institutions. The present ambiguity of terminology is hindering the advance of research and practice in digital libraries and in our ability to communicate the scope and significance of our work.

Keywords: digital libraries, definition, terminology, digital libraries initiatives, information infrastructure, research funding, international, community development, libraries, services, institutions, universities, social aspects

INTRODUCTION

Scholarly and professional interest in digital libraries has grown rapidly throughout the 1990s. In the United States, digital libraries (DL) were designated a national challenge application area under the High Performance Computing and Communications Initiative (HPCC), and a key component of the National Information Infrastructure (Office of Science and Technology Policy, 1994). The Digital Library Initiative (1994-98) involved three U.S. federal agencies. The Digital Libraries Initiative - Phase II (1998-2003) involved eight agencies, indicating the expansion of interest and scope over this short period of time. An international digital libraries program was recently announced by the National Science Foundation, extending the range of partnerships. The United Kingdom has the Electronic Libraries Programme (eLib) (<http://ukoln.bath.ac.uk/elib/>) and many DL research projects are under way in Europe, Asia, and elsewhere, whether under DL-specific funding initiatives or funding from other areas.

During this time period, multiple domestic and international digital libraries conferences were established and digital libraries topics were introduced at meetings in a variety of disciplines and professions. Several new print and online journals on DLs were founded. Online distribution lists with news of DL projects proliferate. Libraries are undertaking projects in digital imaging, document management, and network services.

Why all of this interest and activity? Did an urgent research and development problem lead to large amounts of grant funding? Did the availability of grant funding create opportunities for a new research area? Did successful research lead to practical developments? Did practical problems lead to research on solutions? Is digital library research and practice a definable area of interest, or has "digital library" merely become an umbrella term for a wide array of information and technology projects?

Causal relationships are notoriously difficult to establish. At the rate that the trees of digital library research and practice currently are growing, it is difficult to grasp the shape and size of the forest. We expect the answers to these questions to become ~~more~~ clearer in hindsight, a few years from now. Yet actions we take now, and perceptions that we form, may influence the shape of that forest profoundly. A special issue on digital libraries offers an appropriate venue to take stock of where we are. This article describes the emergence of communities of digital library research and practice, explores definitions of digital libraries arising from these communities, and speculates on future directions.

This is an opinion piece by an active player in U.S. and international DL research activities, intended to provoke discussion in these emerging communities. The opinions expressed are my own and should not be attributed to any of the organizations or collaborators with whom I have worked. In addition to the usual literature review, the paper draws upon related work published elsewhere, including my forthcoming book (Borgman, in press), a history of library automation (Borgman, 1997), and the report of a U.S. National Science Foundation workshop on Social Aspects of Digital Libraries (Borgman et al., 1996) that we conducted at UCLA.

PERSPECTIVES ON DIGITAL LIBRARIES

In a few short years of research and development, already the term "digital library" is used to describe a variety of entities and concepts. Definitions abound (Bishop & Star, 1996; Fox, 1993; Fox, Akscyn, Furuta, & Leggett, 1995; Greenberg, 1998; Lesk, 1997; Levy & Marshall, 1995; Lucier, 1995; Lyman, 1996; Lynch & Garcia-Molina, 1995; Waters, 1998a; Zhao & Ramsden, 1995). A review of these definitions indicates that in general, researchers focus on digital libraries as content collected on behalf of user communities, while librarians focus on digital libraries as institutions or services. These communities are not mutually exclusive. Some

researchers are focusing on practical problems related to institutions and services and some practitioners are participating in research teams addressing issues of content, collections, and communities. In this section we examine possible explanations for these contrasting perspectives. In the following section we return to specifics of digital library definitions.

Research vs. Practice

Despite building upon a foundation of decades of research and practice in related areas, the term "digital library" is relatively new. The availability of research funding under this term has attracted scholars and practitioners from a variety of backgrounds, some of whom have minimal prior knowledge of related areas such as information retrieval, computer networks, cataloging and classification, library automation, archives, or publishing. Sometimes other research topics were simply relabeled "digital libraries," adding to the confusion. The rapid growth in computing networks, databases, and public awareness have contributed to a bandwagon effect in hot topics such as digital libraries, digital archives, and electronic publishing. Only as an area matures do people give serious thought to rigorous definitions.

One reason for the confusion of terminology is that research and practice in digital libraries are being conducted concurrently at each stage of the continuum from basic research to implementation. Some people are working on fundamental enabling technologies and theoretical problems, others are working on applications, others are studying social aspects of digital libraries in experimental and field contexts, and yet others are deploying the results of earlier research. Their concerns and foci are understandably different.

The variety of concerns within the digital libraries research community reflects the interdisciplinary nature of the topic. Scholars based in computer science are largely concerned with enabling technologies and networks. Scholars based in library and information science are largely

concerned with content, organization, user behavior, and publishing. Those based in sociology or economics are more likely to concern themselves with social context and economic models, respectively. Topics such as human-computer interaction, interface design, and service delivery often cross all of these disciplines and more. Scholars based in application areas such as education, geography, health, or arts and humanities may combine any of these areas with expertise in their problem domain. Many, if not most, digital libraries projects draw upon the expertise and research results of multiple disciplines.

Research and practice have a symbiotic relationship. Interesting research problems often arise from practice. Scholars attempt to isolate problems for research purposes and then provide solutions to practitioners for implementation. Partnerships between researchers and practitioners are fundamental to the design of current funding initiatives, encouraging such relationships. Universities are ever more eager to establish partnerships between scholars and industry, in hopes of cross-fertilizing ideas and acquiring new funding sources.

Definitions can serve many purposes, one of which is to provide a focal point for a community. Research-oriented definitions are intended to highlight significant research problems and the relationships between them. They also are intended to attract other scholars, with the goal of achieving a critical mass of researchers to address a given set of problems. Such definitions are more useful in directing attention to research problems than for drawing explicit boundaries.

Practice-oriented definitions are intended to highlight current and anticipated practical challenges. Research libraries in particular are focusing attention on the changing nature of the university, the evolution of libraries as institutions, the role that libraries play in serving the university community, and how that role is changing with the advent of digital collections and services. Librarians are faced with formulating visions for the future of their institutions and services while managing daily operations that may serve tens of thousands of users. The Digital Library

Federation is a consortium of major research libraries whose purpose is to draw attention to these challenges. Their working definition (Waters, 1998a) is a means of framing the practical, rather than research, problems they face as a community.

Community Building Efforts

Researchers and practitioners alike are engaged in deliberate efforts to build communities of interest around digital library issues. Funding agencies are building communities to address digital libraries research problems. They do so through workshops that bring together current and prospective grantees to identify research problems. The first digital libraries initiative also held semi-annual meetings of all the funded research teams, inviting selected observers from the research and practice communities as well. Proposals and collaborative efforts often evolve directly or indirectly from such meetings. Similarly, the Digital Library Federation is an intentional effort to build a community of librarians around practical issues. They do so by collaborating on projects, sharing expertise, and publicizing their initiatives and accomplishments.

The many digital libraries conferences play important roles in community building, as do journals and online news services. Early conferences were organized by individual universities and agencies. Later conference series were supported by professional societies, most notably the Association for Computing Machinery Digital Libraries conferences (Fox, & Marchionini, 1996) and the Advances in Digital Libraries conferences, first supported by the IEEE in 1998 (Proceedings of the IEEE, 1998). Sessions and special interest groups [\[I'm not sure what this phrase means here – eaf\]](#) at related conferences assist in forming and extending communities of research and practice.

Framing the Issues

Digital libraries are attracting interest in many disciplines and professions. While increased participation leads to the cross-fertilization of ideas, it also results in disputed territory and terminology. Clifford Lynch (1993) was prescient in noting that the term "digital library" is problematic because it obscures the complex relationship between electronic information collections and libraries as institutions. Greenberg (1998, p. 106) comments that "the term "digital library" may even be an oxymoron: that is, if a library is a library, it is not digital; if a library is digital, it is not a library." Battin (1998, p. 276-277) rejects the use of the term "digital library" on the grounds it is "dangerously misleading."

As is evident from the definitions presented in the next section, librarians tend to take a broad view of the concept of a "library." Stated in general terms, they see libraries as organizations that select, collect, organize, conserve, preserve, and provide access to information on behalf of a community of users. Libraries have existed for many centuries and their social role and practices have evolved through many forms of civilization and many formats of media. With the advent of computer networks and digital media, libraries will employ yet another delivery system for yet another form of media. In this sense, the term "digital library" connotes "the future library," in which the institution is transformed to address the new environment in which it exists. A sense of continuity and the maintenance of information resources over time ("conserve, preserve") is implicit as well.

Most of the definitions arising from the research community, especially those set forth by computer scientists, tend toward a narrower view of the concept of a "library." Their emphasis is on databases and information retrieval and thus on collecting, organizing, and providing access to information resources. Much of the richer social and institutional context, services, and conservatorship are outside the scope of research-oriented definitions of digital libraries. The narrow scope of the term "library" follows from earlier uses in computer science research and

practice in reference to any collection of similar materials. Rooms housing magnetic tapes are referred to as "tape libraries" and the clerks that check tapes in and out are referred to as "tape librarians," much to the dismay of professional librarians who hold graduate degrees in the field.

The term "digital library" serves as a convenient and familiar shorthand to refer to electronic collections, and conveys a sense of richer content and fuller capabilities than do terms such as "database" or "information retrieval system." At the same time, such uses of the term convey a far narrower sense of a library than one of a full-service institution with long-term responsibilities. Predictions by computer scientists of a declining role for librarians in a digital age (e.g., Odlyzko, 1995, 1997; Schatz, 1997) are predicated on a constrained view of the present and future role of libraries.

Despite the tensions between these perspectives, the communities have not engaged in direct discussion to the extent that might be expected. While the U.S. digital libraries initiatives has shaped the direction of research activities, the library community has made little mention of them or of their influence on conceptions of library services. A salient example is the widely-cited Books, Bricks, and Bytes: Libraries in the Twenty-First Century, first published as a special issue of *Daedalus* (Books, Bricks, and Bytes, 1996) and re-issued as a monograph (Graubard & LeClerc, 1998). The only mention of the digital libraries initiatives is in a piece by the director of the German National Library (Lehmann, 1996); Keller (1998) comments on this point as well. Similarly, only one mention of the digital libraries initiatives can be found in a significant new book on academic information resources for the 21st century published by the Council on Library and Information Resources and the Association of American Universities (Hawkins & Battin, 1998). That mention is by Waters (1998b), the head of the Digital Library Federation.

On the research front, some in library and information science (LIS) take computer scientists to task for reinventing their research on organization of information, information retrieval, user

interfaces, and related topics; they are more likely to do so in conference discussion sessions or in private than in print, however. Computer science researchers counter that LIS researchers are bound by a narrow paradigm and pay insufficient attention to computer science accomplishments. Such sniping increases tensions and is counterproductive to achieving common goals.

Encouraging signs of cooperation and engagement are evident as well. Digital libraries conferences, while dominated by researchers, are drawing contributions and attendance from the practitioner community as well. The diversity of meaning of the term "digital library" continues to be evident in conference programs, however, with odd juxtapositions of papers that bear more similarity in title than in content.

Another sign of cooperation is the emergence of international digital libraries meetings and workshops that draw a balanced audience of researchers and practitioners. A good example is the recent First Asia Digital Libraries Workshop (Yen & Yang, 1998). The organizers were explicit in their intent to attract an emerging community to the event and then to address the mutual interests of those involved. They devoted conference time to the ways and means of developing a community around digital library issues in the Asia-Pacific region. An international conference with similar goals is planned for the spring of 1999 in Dubrovnik, Croatia, entitled CoLIS 3: -Third International Conference on Conceptions of Library and Information Science, with the theme of "Digital libraries: -Interdisciplinary concepts, challenges, and opportunities."

The most significant cooperation may come in the next rounds of research funding for digital libraries. As the scope of these programs expands, expertise and resources resident in the library community will be essential to successful research. Similarly, expertise in networks, enabling technologies, and related areas resident in the computer science community will be essential as

well. Hopefully, increased collaboration will enhance mutual respect for and learning about each other's fields.

DEFINING DIGITAL LIBRARIES

We return to the development of definitions of the term "digital library" arising from the research and practice communities. First we explore the origin and evolution of research-oriented definitions, then examine definitions set forth by library practitioners. Lastly we consider how these definitions might be applied to the array of extant electronic databases, many of which are referred to as "digital libraries."

Digital Libraries as Content, Collections, and Communities

Digital library research builds upon a long history of related work in information retrieval, databases, user interfaces, networks, information seeking, classification and organization, library automation, publishing, and other areas. It dates back several decades or centuries, depending on what is included for consideration. We include in the research community scholars studying information-related problems that they or others have labeled "digital libraries." Most of these scholars are affiliated with academic departments or research groups in computer science, library and information science, or information studies, but some are located in related areas such as sociology, psychology, communication, or economics, or in application areas such as education, geography, health sciences, or the arts and humanities.

Definitions of digital libraries arising from the computer and information science research community have evolved in scope and content throughout the 1990s. The two initiatives funded by the multiple U.S. federal agencies (National Science Foundation, 1993, 1998) [1] were particularly influential in defining the boundaries of digital libraries research. The definitions

were not established by the funding agencies alone. Rather, they arose from the many research workshops and conferences that took place before and during the initiatives, as well as from publications by researchers.

Research-oriented Definitions

One of the first meetings to focus directly on digital libraries issues was a 1991 workshop on "Future Directions in Text Analysis, Retrieval and Understanding" (summarized in Fox, 1993) and a white paper on electronic libraries that followed from it (Lesk, Fox, & McGill, 1991). The resulting excitement led to more workshops that refined a research agenda and eventually led to the digital libraries initiatives.

The earliest research-oriented definition appears to be one that I proposed in 1992 for what were then called "electronic libraries." It was included in a sourcebook of materials for those preparing proposals to the Digital Library Initiative (Fox, 1993). Summarizing from workshops conducted in 1991 and 1992, this definition states that a National Electronic Library is (1) a service; (2) an architecture; (3) a set of information resources, databases of text, numbers, graphics, sound, video, etc.; and (4) a set of tools and capabilities to locate, retrieve, and utilize the information resources available. The users of a national electronic library would include students, teachers/professors, researchers/scholars, librarians, authors, publishers, information providers, and practitioners. Contributors of information resources would include publishers, universities, professional societies, libraries, authors, editors, and compilers.

The above definition remains among the most comprehensive by including services, architecture, content, enabling technologies, users, and content. It provided a basis for further discussion and refinement. The Digital Library Initiative, announced in September, 1993 (and since dubbed "DLI-1"), defined the term only implicitly (National Science Foundation, 1993), stating that

Information sources accessed via the Internet are the ingredients of a digital library. Further, the problem for research and development is ... to achieve an economically feasible capability to digitize massive corpora of extant and new information from heterogeneous and distributed sources; then store, search, process and retrieve information from them in a user friendly way. Note the use of the singular form, "digital library," having evolved from the goal of "a national electronic library."

The goals of DLI-1 were modest by today's standards. Research was supported in three areas: (1) capturing data and metadata of all forms (text, images, sound, speech, etc.) and categorizing and organizing them, (2) advanced software and algorithms for browsing, searching, filtering, abstracting, summarizing and combining large volumes of data, imagery, and all kinds of information; and (3) the utilization of networked databases distributed around the nation and around the world.

The notion of a digital library and goals for research continue to be refined through workshops, conferences, and scholarly writing. A 1995 NSF [Actually the major responsibility was with ARPA, I believe, so it is more terse to not give a sponsor.] workshop that addressed scaling and interoperability issues in digital libraries resulted in several definitions, the most general of which defines a digital library as a system that provides "a community of users with coherent access to a large, organized repository of information and knowledge" (Lynch & Garcia-Molina, 1995). Content, collection, and community all are included in this definition, as well as the requirement that the content be organized.

One of the primary outcomes of the NSF-sponsored Social Aspects of Digital Libraries workshop was a definition of the term "digital libraries." We broadened the scope to encompass two complementary ideas (Borgman, et al., 1996):

1. Digital libraries are a set of electronic resources and associated technical capabilities for creating, searching, and using information. In this sense they are an extension and enhancement of information storage and retrieval systems that manipulate digital data in any medium (text, images, sounds; static or dynamic images) and exist in distributed networks. The content of digital libraries 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.

2. Digital libraries are constructed Ñ collected and organized Ñ by [and for] 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 above definition extends the scope of digital libraries in several directions, reflecting the contributions of scholars from a dozen disciplines. It moves beyond information retrieval to include the full life cycle of creating, searching, and using information. Rather than simply collecting content on behalf of user communities, it embeds digital libraries in the activities of those communities, and it encompasses information-related activities of multiple information institutions.

Like the first initiative, the call for proposals for the Digital Libraries Initiative - Phase II

(National Science Foundation, 1998) does not include an explicit definition of the term "digital library." Rather, the call notes that since the first initiative, "the definition of a digital library has evolved." Compared to the first initiative, the DLI-2 call includes far more concern for social, behavioral, and economic aspects of digital libraries, and ~~includes~~ identifies research areas that encompass a broader range of academic disciplines, reflecting most of the issues raised in the definition from the Social Aspects of Digital Libraries workshop. Most of the call focuses on research, which is divided into (1) human-centered research, (2) content and collections-based research, and (3) systems-centered research. The remainder addresses testbeds and applications. Explicit in the call is a view of digital libraries as a component of a national and international information infrastructure. A sense of service to user communities is implicit in these new directions as well.

Other definitions require that digital libraries contain the full content of information resources in computer-readable form, and often assume that they contain more than text alone (Fox, et al., 1995; Levy & Marshall, 1995). Lesk (1997), in a book on "practical digital libraries," defines a DL simply as "a collection of information which is both digitized and organized." Summarizing a broad array of DL definitions, Bishop and Star (1996) determined that three elements are necessary: (1) some sense of a collection, with some kind of organization; the content may be partly physical and partly electronic, or entirely electronic; (2) a collection that is not entirely bibliographic or exclusively a set of pointers to other material -- it must contain some "full-form online material" and may be in a variety of formats; and (3) a goal exists to link "audience, group, patron, or community with attributes of the collection," whether in the manner that physical collections are selected for an audience or in the sense of the virtual space that can be created around a community.

Defining Elements of Digital Libraries

Several aspects of these definitions should be noted. One is that digital libraries are viewed as databases, albeit databases of rich content, whether full text, images, or combinations of media and representations. Much digital library research, particularly that conducted in departments of computer science, focuses on "enabling technologies" such as database structure, retrieval algorithms, filtering, intelligent agents, network architecture, and other necessary capabilities.

These definitions assume or require that content is collected on behalf of a user community. This aspect of the definition frames digital libraries in terms of their users, which also determines the tools and capabilities those users need to manipulate the content. Digital library research on information needs and uses, users, interface design, and social context derives from these aspects of the definitions. The one definition that mentions institutions indicates that digital libraries can be extensions of libraries, museums, archives, and schools, as well as extensions of work, education, and leisure environments in which information resources are used. The notion of "community" remains problematic, as none of these definitions provide criteria for identifying or determining the scope of a user community.

Another noteworthy assumption, particularly in definitions originating in the U.S., is that digital libraries exist in distributed environments. This is not surprising, given that the U.S. digital libraries initiatives are closely related to information infrastructure development (National Science and Technology Council, 1998; Office of Science and Technology Policy, 1994). The DLI-1 (National Science Foundation, 1993) call for proposals begins by defining the Internet and setting the need for DL research in a network context. By the time DLI-2 was announced five years later (National Science Foundation, 1998), technical issues of operating digital libraries on computer networks had become core research concerns. These include interoperability, portability, data exchange, scalability, federation, extensibility, and open network architectures.

Digital Libraries as Institutions or Services

The terms "digital library," "electronic library," and "virtual library" have appeared in the professional literature of library and information science for some years already, but rarely with explicit definitions. Lyman (1996), in an article entitled "What is a digital library? Technology, intellectual property, and the public interest," explores concepts he views to be prerequisite to defining the concept of a digital library, such as electronic publishing and digital documents. Young (1996) lists characteristics of digital libraries, saying they "provide personalized or custom services for accessing, assembling, and analyzing information resources from a variety of diverse sources in many different formats" (1996, p 122).

Waters (1998a) provides the first succinct definition from a librarian's perspective. This is the working definition set forth by the Digital Library Federation (DLF):

Digital Libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.

Waters acknowledges that the DLF statement is broad and is intended to comprehend other uses of the term, including those of the digital libraries initiatives. However, it is distinct from the research-oriented definitions in several respects. The focus of the DLF definition is on the digital library as an organization whose services include the provision of information resources in digital forms. The institutional focus also adds the element of conservatorship in preserving integrity and ensuring persistence of digital collections. The DLF definition thus captures a much broader sense of the term "library."

The DLF definition captures the senses and sentiments implicit in descriptions of digital libraries by other library agencies. The Library of Congress's American Memory project, for example, is part of a national digital library. Note the singular form "digital library" in reference to an institution that provides many individual resources. American Memory is a project to create digital collections of historical photographs and related materials and providing public access via the Internet (<http://memory.loc.gov>). It is similar to projects undertaken by other members of the DLF (<http://www.clir.org/programs/diglib/diglib.html>).

The California Digital Library (CDL) is a tenth library for the University of California, which has nine campuses (<http://www.cdlib.org>). Again we find the sense of an institution that offers services: "it is responsible for the design, creation, and implementation of systems that support the shared collections of the University of California."

Funding for digital library projects in the United Kingdom is built upon a framework more similar to the DLF definition than to definitions arising from the U.S. digital libraries initiatives. Primary funding in the U.K. is provided by the Electronic Libraries (eLib) Program, which is a program of collaborative partnerships between libraries, academic staff, publishers and others, to promote large scale pilot and demonstrator projects focusing on the components of a future electronic library service, and to provide stimulus to its creation. (Rusbridge, 1998). ELib focuses on developing services rather than on basic research or enabling technologies. Libraries are the primary institutions involved, though eLib promotes partnerships with publishers, academic staff (university faculty in American parlance), and others.

Similarly, the summary report of the British Library's Initiatives for Access Programme, entitled Towards the Digital Library, suggests a definition similar to that of the DLF (Carpenter, Shaw, & Prescott, 1998). Mahoney, in an overview chapter, says: "We were clear from the beginning that for The British Library, and probably for most libraries, being a digital library

would emphatically not mean being an exclusively digital library (p. 11). Later, Mahoney implies that a digital library is a service: "the issue of how one moves ... towards a more integrated form of digital library provision" (p. 17).

The above statement from the British Library highlights another problem with using the term "digital library" to refer to an institution. Libraries collect content based on the information needs of their user communities; the medium in which the content is captured is a secondary concern. By defining an institution in terms of the format of materials (digital content), then we risk distinguishing between print libraries, digital libraries, film libraries, audio libraries, and so on. Librarians are not entirely comfortable with the term "digital library" to describe some future form of the institution, but have yet to propose a suitable alternative.

Digital Libraries as Databases

Neither the definitions arising from research nor practice communities deal explicitly with the plethora of databases that exist on the Internet, the World-Wide Web, on CD-ROMs, and on proprietary services such as Dialog, Lexis/Nexis, Westlaw, STN, InfoAmerica, and CDB Infotek. Some of these databases and web sites identify themselves as digital libraries, whether for reasons of scholarship, for convenience as a recognizable term, or as a marketing ploy. In other cases, surveys of digital libraries include web-based, CD-ROM, and other databases within their scope.

These databases fall into a grey area between the definitions constructed by the research and library communities. The lack of fit is not surprising, as neither definition was intended to categorize electronic databases. We can say that electronic databases per se are not libraries as institutions or services, in the sense of the DLF definition. Griffiths (1998) confronts the question of "why the web is not a library." Her reasons include incompleteness of content, lack

of standards and validation, minimal cataloging, and ineffective information retrieval. To this we add that the World-Wide Web is not an institution and is not organized on behalf of a specifiable user community. However, one of the services that digital libraries, in the DLF sense, provide is access to electronic databases.

Some portion of electronic databases on the Internet, on proprietary systems, and on CD-ROMs are digital libraries in the senses defined by the research community. On a case-by-case basis we can judge the degree to which given databases are organized collections, whether they were created for a specified community, and whether their capabilities are sufficient to distinguish them from other forms of information retrieval systems, for example.

INTO THE DIGITAL FUTURE

Despite the progress made to date, we are still in the early stages of digital libraries research and practice, under any of these definitions. Where do we go from here? Many signs point to ever greater developments and investments in networked information technologies (Lynch, 1998). In the United States, the successful Congressionally-chartered High Performance Computing and Communications program (HPCC) is succeeded by the Executive Branch's Computing, Information, and Communications (CIC) research and development program (National Science and Technology Council, 1998). Research areas include high end computing and computation, large scale networking, including the Next Generation Internet Initiative, high confidence systems, human centered systems, and education, training, and human resources. Digital libraries research now falls under the human-centered systems program of CIC.

The U.S. is not alone in promoting the development of information and communication technologies. The Group of Seven [2] major industrialized nations (G-7) supports the

development of a global information infrastructure (G-7 Ministerial Conference on the Information Society, 1995a, b). One of the G-7 policy statements is that the GII will provide access to culturally and linguistically diverse content, and thus involves digital libraries. The European Union (EU) funds and promotes a wide range of information-related research and development under Directorate-General XIII, Telecommunications, Information Market and Exploitation of Research. Many other countries have established national information infrastructure programs and associated research and development support mechanisms.

The continued expansion of information infrastructure and the penetration of information technology into more aspects of daily activities will require basic and applied research in many disciplines. Just as the frontiers of computing have moved from desktop to mobile computing to embedded systems, digital libraries are themselves becoming "enabling technologies" for many other applications. Contributing materials to DLs is a form of electronic publishing, and materials published electronically are collected, organized, preserved, and disseminated electronically. Distance-independent learning requires that content be associated with instruction, hence DLs are an essential component. Software that supports computer-supported cooperative work must include a means to manage the associated work products, which is a digital library problem. And so on.

Many fundamental technical problems in digital libraries research remain to be solved. As digital libraries become more sophisticated, more practical, and more embedded in other applications, the challenges of understanding their uses and users become ever more urgent. These are inherently interdisciplinary problems, and will require the contribution of researchers from many backgrounds. Some of them have yet to hear the term "digital libraries," much less recognize that their interests are relevant.

Digital libraries as institutions or services stand to benefit from research on almost all aspects of

digital libraries as content, collections, and communities. Research libraries and universities engaged in reinventing themselves for a digital age will need to draw upon the best research, theory, and practice from a myriad of disciplines. These are urgent challenges. As Hawkins (1998) puts it, the traditional library is unsustainable in its present form.

Conversely, researchers studying many digital library problems will need partnerships with library institutions to study and test in operational settings. This is especially true of research on social, behavioral, and economic aspects of digital libraries. Partnerships with other information institutions such as archives, museums, and schools will be essential as well.

SUMMARY AND CONCLUSIONS

Interest in digital libraries research and practice has expanded rapidly throughout the 1990s. Major funding initiatives in the U.S., U.K., European Union, and elsewhere have fueled research and development. Conferences, journals, and news services on digital libraries proliferate. Upon closer examination, we find that the term "digital library" has multiple meanings. These definitions cluster around two themes. From a research perspective, digital libraries are content collected and organized on behalf of user communities. From a library practice perspective, digital libraries are institutions or organizations that provide information services in digital forms.

Definitions are formulated to serve specific purposes, and such categorization is not the purpose of the definitions identified in this article. The research community's definitions serve to identify and focus attention on research problems and to expand the community of interest around those problems. The library community's definitions focus on practical challenges involved in transforming library institutions and services. Hence neither the research community nor library community definitions are particularly helpful in categorizing the vast array of

databases available on the Internet, on proprietary services, and on CD-ROMs.

Given the rapid expansion of computer networks, distributed access to information resources, electronic publishing, distance-independent learning, electronic commerce, and related technologies, vastly more research on all aspects of digital libraries is needed. Technological developments, structural changes in the way universities are funded, escalating costs of information resources, demographics, and other factors combine to make the transformation of research libraries an urgent challenge. Attention to digital libraries research and practice, under all of these definitions, will continue to increase for the foreseeable future.

In exploring these definitions, we find that the research community's definitions have evolved from a narrower view emphasizing enabling technologies to one that encompasses the social, behavioral, and economic contexts in which digital libraries are used. That view also has expanded from a primary emphasis on information retrieval to include the full life cycle of creating, seeking, using, preserving, and disposing of information resources. The library community has voiced the term "digital library" for some years, but only recently has promulgated formal definitions. The working definition set forth by the Digital Libraries Federation appears to capture the senses in which practicing librarians intend the term.

At present, the term "digital library" is being used with two distinctly different meanings. Taken together, the two definitions result in a tautology: A digital library is an institution that provides digital libraries. Both definitions are problematic because they confuse the boundaries between electronic collections and institutions. Underlying most of the research-oriented definitions is a constrained view of the nature of libraries. Yet using the term to imply the broader view favored by librarians constrains the institution by the type of content it collects.

Neither community is likely to surrender the term in favor of another. Given this inherent

conflict of interest, people using the term need to define what they mean in context. The failure to define terms slows the development of theory, research, and practice. It also limits the ability to communicate the scope of the area or the nature of the research and practice problems to others. While all parties need not agree on one meaning, each can be more explicit in explaining choices of terminology. Sometimes we simply may need to agree to disagree. Words do matter, and will influence the success of our ventures. I hope this article will stimulate discussion of what digital libraries are and what they can be in the future.

ACKNOWLEDGEMENTS

My thinking about digital libraries has evolved over many years through discussions with far more people than can be acknowledged here. Particularly influential were the many workshops and meetings related to the Digital Libraries Initiatives, the Knowledge and Distributed Intelligence Initiative, the Social Aspects of Digital Libraries workshop we conducted at UCLA, the ACM Digital Libraries and Information Retrieval conferences, and meetings of the board of the Council on Library Resources and later the Council on Library and Information Resources. This article is an outgrowth of my forthcoming book (Borgman, in press). Commentary and suggestions by Philip Agre (UCLA) on definitional aspects of the book contributed to shaping the ideas in this article, as did comments by Tefko Saracevic on a preliminary draft submitted to COLIS 3, and comments by Gary Marchionini and Edward Fox on an earlier draft of the article. Students in my fall, 1998, PhD Seminar on Information Seeking and Use discussed an earlier draft and provided critical commentary. The final draft was refined further through a presentation to the UCLA Information Studies Seminar series coordinated by Prof. Agre. Many thanks are due to my students and colleagues for exploring these issues with me. I accept credit and blame for the opinions, facts, and interpretations presented here.

NOTES

[1]. The Digital Library Initiative (1994-1998) was jointly funded by the U.S. National Science Foundation, Computer and Information Science and Engineering Directorate; the Advanced Research Projects Agency Computing Systems Technology Office and the Software and Intelligent Systems Technology Office; and the National Aeronautics and Space Administration. The Digital Libraries Initiative - Phase II (1998-2003) is funded by the above three agencies plus the National Library of Medicine, the Library of Congress, and the National Endowment for the Humanities, in partnership with the National Archives and Records Administration and the Smithsonian Institution.

[2]. The Group of Seven nations are Canada, France, Germany, Italy, Japan, the United States, and the United Kingdom.

REFERENCES

Battin, P. (1998). Leadership in a transformational age. In B.L. Hawkins & P. Battin (eds.), The Mirage of Continuity: Reconfiguring Academic Information Resources for the 21st Century. Washington, DC: Council on Library and Information Resources and the Association of American Universities. pp. 260-270.

Battin, P., & Hawkins, B. L. (1998). Setting the stage: -Evolution, revolution, or collapse? In B.L. Hawkins & P. Battin (eds.), The Mirage of Continuity: Reconfiguring Academic Information Resources for the 21st Century. Washington, DC: -Council on Library and Information Resources and the Association of American Universities. pp. 3-12.

Bishop, A.P.; & Star, S.L. (1996). Social informatics for digital library use and infrastructure. In. M.E. Williams (ed.), Annual Review of Information Science and Technology, 31. Medford,

NJ; Information Today, pp. 301-401.

Books, Bricks, and Bytes. (1996). Daedalus, Journal of the American Academy of Arts and Sciences; Proceedings of the American Academy of Arts and Sciences, 125(4).

Borgman, C. L. (in press). From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked World. Cambridge, MA: MIT Press.

Borgman, C.L. (1997). From acting locally to thinking globally: A brief history of library automation. Library Quarterly, 67(3), 215-249.

Borgman, C.L.; Bates, M.J.; Cloonan, M.V.; Efthimiadis, E.N.; Gilliland-Swetland, A.; Kafai, Y.; Leazer, G.L.; Maddox, A. (1996). Social Aspects Of Digital Libraries. Final report to the National Science Foundation. [Hhttp://dli.grainger.uiuc.edu/national.htm](http://dli.grainger.uiuc.edu/national.htm)

Carpenter, L.; Shaw, S.; & Prescott, A. (eds.). (1998). Towards the Digital Library: The British Library's Initiatives for Access Programme. London: The British Library.

Fox, E.A. (ed.). (1993). Sourcebook on Digital Libraries: Report for the National Science Foundation, TR-93-35, VPI&SU Computer Science Department, December, 1993, Blacksburg, VA, 439 pages. Available by anonymous FTP from directory pub/DigitalLibrary on fox.cs.vt.edu or at <http://fox.cs.vt.edu/DLSB.html>.

Fox, E. A., Akscyn, R. M., Furuta, R. K., & Leggett, J. J. (1995). Digital Libraries [Special issue [introduction](#)]. Communications of the ACM, 38(4), 22-28.

Fox, E.A.; & Marchionini, G. (eds.). (1996). Proceedings of the 1st ACM International

Conference on Digital Libraries. New York: Association for Computing Machinery.

G-7 Ministerial Conference on the Information Society. (1995a). Chair's conclusions.
<http://www.ispo.cec.be/g7/g7main.html>.

G-7 Ministerial Conference on the Information Society. (1995b). Electronic libraries project.
<http://www.ispo.cec.be/g7/g7main.html> or <http://www.cec.lu>.

Graubard, S.R. & LeClerc, P. (eds.). (1998). Books, Bricks, and Bytes: Libraries in the Twenty-First Century. New Brunswick, NJ: Transaction Publishers.

Greenberg, D. (1998) Camel drivers and gatecrashers: Quality control in the digital research library. In B.L. Hawkins & P. Battin (eds.), The Mirage of Continuity: Reconfiguring Academic Information Resources for the 21st Century. Washington, DC: Council on Library and Information Resources and the Association of American Universities. pp. 105-116.

Griffiths, J-M. (1998). Why the web is not a library. In B.L. Hawkins & P. Battin (eds.), The Mirage of Continuity: Reconfiguring Academic Information Resources for the 21st Century. Washington, DC: Council on Library and Information Resources and the Association of American Universities. pp. 229-246.

Hawkins, B. L. (1998). The unsustainability of the traditional library and the threat to higher education. In B.L. Hawkins & P. Battin (eds.), The Mirage of Continuity: Reconfiguring Academic Information Resources for the 21st Century. Washington, DC: Council on Library and Information Resources and the Association of American Universities. pp. 129-153.

Hawkins, B.L., & Battin, P. (eds.) (1998). The Mirage of Continuity: Reconfiguring Academic

Information Resources for the 21st Century. Washington, DC: -Council on Library and Information Resources and the Association of American Universities.

Keller, M.A. (1998). Libraries in the digital future. Science, 281, 1461-1462.

Lehmann, K-D. (1996). Making the transitory permanent: -The intellectual heritage in a digitized world of knowledge. In Books, Bricks, and Bytes, Daedalus, Journal of the American Academy of Arts and Sciences; Proceedings of the American Academy of Arts and Sciences, 125(4), 307-330. Republished in S.R. Graubard & P. LeClerc (eds.). (1998). Books, Bricks, and Bytes: Libraries in the Twenty-First Century. New Brunswick, NJ: -Transaction Publishers.

Lesk, M.E. (1997). Practical Digital Libraries: Books, Bytes, and Bucks. San Francisco: Morgan Kaufman.

Lesk, M.; Fox, E.; & McGill, M. (eds.) (1991). A National Electronic Science, Engineering, and Technology Library. Reprinted in Fox (1993), pp. 4-24.

Levy, D. M.; & Marshall, C. C. (1995). Going Digital: A look at the assumptions underlying digital libraries. Communications of the ACM, 38(4), 77-84.

Lucier, R.E. (1995). Building a digital library for the health sciences: Information space complementing information place. Bulletin of the Medical Library Association, 83(3), 346-350.

Lyman, P. (1996). What is a digital library? Technology, intellectual property, and the public interest. In Books, Bricks, and Bytes, Daedalus, Journal of the American Academy of Arts and Sciences; Proceedings of the American Academy of Arts and Sciences, 125(4), 1-33.

Republished in S.R. Graubard & P. LeClerc (eds.). (1998). Books, Bricks, and Bytes: Libraries in the Twenty-First Century. New Brunswick, NJ: -Transaction Publishers.

Lynch, C.A. (1993). Accessibility and Integrity of Networked Information Collections (Background Paper No. BP-TCT-109). Washington: Office of Technology Assessment.

Lynch, C. (1998). The evolving Internet: -Applications and network service infrastructure. Journal of the American Society for Information Science, 49(11), 961-972.

Lynch, C.; Garcia-Molina, H. (1995). Interoperability, scaling, and the digital libraries research agenda. <http://www.hpcc.gov/reports/reports-nco/iita-dlw/main.html>.

National Science Foundation. (1993). Research On Digital Libraries, NSF 93-141, A Joint Initiative of: National Science Foundation Computer and Information Science and Engineering Directorate; Advanced Research Projects Agency, Computing Systems Technology Office and the Software And Intelligent Systems Technology Office; and the National Aeronautics and Space Administration. Washington, DC: -National Science Foundation.

National Science Foundation. (1998). Digital Libraries Initiative - Phase II. Call for Proposals. Arlington, VA: -National Science Foundation.

<http://www.nsf.gov/pubs/1998/nsf9863/nsf9863.htm>

National Science and Technology Council; Committee on Technology; Subcommittee on Computing, Information, and Communications R&D. (1998). Networked Computing for the 21st Century. Supplement to the President's FY 1999 Budget. Washington, DC: Office of Science and Technology Policy.

Odlyzko, A. M. (1995). Tragic loss or good riddance? The impending demise of traditional scholarly journals. International Journal of Human-Computer Studies, 42, 71-122.

Odlyzko, A. M. (1997). Silicon dreams and silicon bricks: the continuing evolution of libraries. Library Trends, 46 (1), 152-167.

Office of Science and Technology Policy; Federal Coordinating Council for Science, Engineering, and Technology; Committee on Physical, Mathematics, and Engineering Sciences. (1994). High Performance Computing and Communications: Toward a National Information Infrastructure. Washington, DC: -Executive Office of the President, Office of Science and Technology Policy.

Proceedings of the IEEE International Forum on Research and Technology Advances in Digital Libraries (ADL Ô98). (1998). April 22-24, Santa Barbara, CA. Los Alamitos, CA: -IEEE Computer Society.

Rusbridge, C. (1998). In Proceedings of the IEEE International Forum on Research and Technology Advances in Digital Libraries (ADL Ô98). (1998). April 22-24, Santa Barbara, CA. Los Alamitos, CA: -IEEE Computer Society.

Schatz, B.R. (1997). Information retrieval in digital libraries:- Bringing search to the Internet. Science, 275(N5298):- 327-334.

Waters, D.J. (1998a). What are digital libraries? CLIR (Council on Library and Information Resources) Issues, No. 4. <http://www.clir.org/pubs/issues/issues04.html>.

Waters, D.J. (1998b). Steps toward a system of digital preservation:- Some technological,

political, and economic considerations. In B.L. Hawkins & P. Battin (eds.), The Mirage of Continuity: Reconfiguring Academic Information Resources for the 21st Century. Washington, DC: Council on Library and Information Resources and the Association of American Universities. pp. 193-206.

Yen, J.; & Yang, C.C. (eds.). (1998). East meets West. Proceedings of the First Asia Digital Library Workshop, August 6-7, Hong Kong. Hong Kong: -University of Hong Kong Libraries, No. 8.

Young, P.R. (1996). Librarianship: A changing profession. In Books, Bricks, and Bytes, Daedalus, Journal of the American Academy of Arts and Sciences; Proceedings of the American Academy of Arts and Sciences, 125(4), 103-125. Republished in S.R. Graubard & P. LeClerc (eds.). (1998). Books, Bricks, and Bytes: Libraries in the Twenty-First Century. New Brunswick, NJ: -Transaction Publishers.

Zhao, D.G.; & Ramsden, A. (1995). Report on the ELINOR electronic library pilot. Information Services & Use, 15, 199-212.