The Role of Digital Libraries in Providing Access to Networked Information*

Farley W. Jenkins†

August 14, 2011

University of South Florida

Abstract

This paper introduces the concepts of a Global Information Infrastructure (GII) and a global digital library. A comprehensive literature review is conducted to explore the role a global digital library might play in providing network access to information within a GII based on current and emerging paradigms in the information sciences and information service professions. Particular attention is given to the concepts of system usability and content usefulness as well as the role of digital librarianship in present and future information service practice. Conclusions are drawn based on the observations of the studies reviewed within the conceptual framework discussed.

†farley.jenkins@gmail.com
1 Introduction

The changes brought about by the advent of graphical browser interfaces to the World Wide Web in 1995 (Tomer, 2010) are truly astonishing. The sheer volume of documents that Web technology enables the creation of and access to has been described as a revolution and has drawn many comparisons with the Gutenberg press (Borgman, 2003). Certainly the comparison with Gutenberg is apt, as one would be hard pressed to think of another event in world history that so dramatically increased the volume of and universality of access to information. The number of books available in Europe increased from about 30,000 on the eve of his invention to over 9 million by the year 1500 (Garibay, Gutiérrez, & Figueroa, 2010). This preponderance of written information as compared to what was available before caused lasting changes in Western society that can accurately be described as revolutionary. Gutenberg’s native Germany transformed from a society with a roughly 50% adult male literacy rate in 1500 (Adler & Pouwels, 2005) to a society with an adult literacy rate of 99% as of the time of this writing (CIA, 2011). The current exponential increase in the availability of electronic information promises to bring changes even more profound and far-reaching.

The information revolution of the 21st century is bringing changes even more sweeping than Gutenberg’s due to two factors. First, the new information resources are networked. While networked machines are certainly more volatile than printed sources of information, they are capable of acting in unison. An author can describe another’s work in a book, but can provide a hyper-reference that allows instant and accurate transfer in a hyper-text page. Their capabilities are mutually reinforcing, thus the current revolution has been characterized by a geometric increase in available information as opposed to the linear increase that marked the Gutenberg revolution. Networked information allows all users at all points of the network to broadcast as well as receive information, resulting in the first truly many-to-many communications network in world history. Second, it is global. While books and periodicals have been exported in the past, the speed of communication depended on available means of transportation, legal boundaries such as tariffs and economic sanctions, as well as existing trading relationships. The World Wide Web allows users to publish with the push of a button, and an audience anywhere in the world with Internet access can view this information within seconds. This is causing the world to feel like a much smaller place, and it has eliminated borders for ideas.

As access to the Web and its myriad of information sources continues to proliferate throughout the globe, the pace of the information revolution is building to
1 INTRODUCTION

a crescendo that Borgman (2003) terms Global Information Infrastructure or GII. GII is not a present reality, but rather an envisioned future state of affairs when the entire world’s population will have access to digital communications technology much like most every home in the United States now has access to electricity. However, the very ease of publishing the building of GII relies on causes the problem of locating trustworthy information sources in a timely manner. As such, digital libraries are beginning to respond to this need by organizing and evaluating electronic information sources in the same manner that physical libraries sort out printed information. The term “digital libraries” has been defined in a number of ways by a variety of sources (Borgman, 2003), but use of such a term connotes the professional cataloging, indexing, and reference services that have long been associated with libraries. A digital library could take the form of an electronic extension of a physical library collection, such as the Library of Congress’ American Memory Project, or a collection of digital documents within a library without walls such as the Internet Archive. As the interconnectivity of digital libraries increases, they are building towards a state of universal access termed the global digital library that Borgman (2003) sees growing in tandem with GII.

As the enabling technologies become more stable and affordable, and as digital libraries become less of a computer science experiment and more of a new method of service provision, an emerging professional group of digital librarians have begun to turn their attention to the role they will play in providing access to information in an environment where all information is available to anyone, anywhere at the push of a button. Borgman (2003) identifies three key components to information access. The first is connectivity. This is a building block of the GII that is falling into place, as the global Internet-using population grew 444.8% from 2000 to 2010 (Miniwatts Marketing Group, 2010). The second is content and services, an area which many studies have labeled usefulness (Tsakonas & Papatheodorou, 2006; Tsakonas & Papatheodorou, 2008; Xie, 2006; Xie, 2008). The third is the usability of information systems and the documents they provide. This aspect has many components, to include ease of use, ease of access, as well as information literacy. Usability is a topic that has received extensive attention in the professional literature, as well as a design goal described as a “moving target” (Borgman, 2003, p. 168). As such, any comprehensive analysis of the role of digital libraries in providing information access within the current and emerging network environments must pay considerable attention to the usability of systems design as well as giving consideration to the usefulness of content.
2 Literature Review

2.1 Usability

Usability analysis has become something of a Rorschach test in digital library design, as one’s perception of an information system’s usability is likely to reveal more about the perspective of the individual than any universally applicable set of design principles. Xie (2008) in particular noted that one size truly does not fit all in the evaluation of digital library usability, as different users will have different needs. Garibay et al. (2010) noted an inclination to rely on numbers in usability evaluation. This tendency towards reductionism is understandable given the complexity of the task, but a myopic focus on metrics is likely to result in a distorted image of user needs. The 5S framework of Gonçalves, Moreira, Fox, and Watson (2007) is an example of such a purely quantitative framework. The first part of the study consisted largely of a series of complex equations that are inaccessible to those without a background in advanced mathematics. The second part consisted of a focus group session with a group of librarians. The librarians raised a number of utility issues in the conversation. These included the usefulness of the figures produced by the formulae in question as well as the applicability of the numbers to the needs of users. While such measures as automated citation counts of journal articles are theoretically tied to needs of users, the lack of user involvement in the evaluation criteria was of concern.

A model of user-centered evaluation was provided by Tsakonas and Papatheodorou (2006). Their evaluation took the form of questionnaires sent to expert users of professional literature databases in chemical engineering and information science. While system performance issues were considered by the study, the investigators noticed that users tend to be most concerned with usefulness, defined as the interaction between user and content, as well as usability, defined as the interaction between user and system features. Participants noted that, if given a choice, they would choose useful content over a usable system, as access to large amounts of authoritative information is critical to their jobs. The biggest bar to usability identified in the study was the sometimes confusing terminology used by the system, which the researchers remarked had the potential to create a “hostile information environment” (Tsakonas & Papatheodorou, 2006, p. 403). One user stated his preference for search engines such as Google over gateway databases, as these helped him identify other useful content, such as conference proceedings and course notes, which were filtered by the professional database. This is a sentiment echoed by Joint (2010), who noted the failure of library federated
search engines to improve upon Google. Rowlands et al. (2008) revealed a very heavy reliance on Web search engines as opposed to library gateways in an extensive study of undergraduate information-seeking behavior and recommended tighter integration of commercial search engines with library Websites in order to increase access for this segment of the academic user population.

The concept that search is central to increasing access to information for expert users is echoed in many studies, particularly in Komlodi, Marchionini, and Soergel (2007). This study looked at the search needs of attorneys and law librarians through observing user interaction with the Westlaw legal database as well as focused interviews. Next, the investigators used what was learned in the first phase of the study to design a search interface for a legal database with the features identified as having the greatest potential to increase information access. Additional observation was conducted and the user interface was further refined in an iterative design process. These features eliminated much of the need for manual note-taking by automatically saving all queries (allowing users to perform a search at a later date to check the current status of legal decisions) and allowing users to add annotations to their search history. Collaboration features allowed users to form working groups and share their searches, complete with their annotations, with their colleagues. Users could search within result sets, combine search results and weed out searches later deemed to be false starts. Aggregated search features allowed users to search across intranets, databases, the Web, and their local hard drives for electronic information stored in any format (to include emails, Web pages, and calendar appointments). Citation chaining support allowed users to follow the history of complex legal decisions. It is difficult to imagine Google or any other search engine providing features that enhance usability more than those resulting from this research.

Tsakonas and Papatheodorou (2008) used a questionnaire format to gauge the satisfaction of users of institutional repositories of open access scholarly articles. The investigators noted that ease of use of advanced features, such as aggregated search functions, played the biggest role in user satisfaction of system usability, but that factors such as layout and aesthetic appearance also played a role. They also noted that users were more satisfied with a system that allowed them to view information at multiple levels, such as viewing an abstract prior to downloading a full-text article. The researchers advocated a unified treatment of usability and usefulness in evaluations. They observed that institutional repositories that allow users to upload their own work had particularly high usefulness satisfaction rates as this model transfers responsibility for content growth to the user.

Xie (2006) developed the concept of user-centered evaluation even further, as
he enlisted users in devising the evaluation criteria. Graduate students of information science were recruited to develop their own questionnaire and then evaluate a number of digital libraries by their own standards. The subjects were also asked to rate individual criteria according to the features most important to them. Most participants identified usability as the most important criteria, although several raised concerns regarding usefulness. For example, more than one subject gave poor marks to digital libraries that did not give information regarding the authority or copyright status of information objects presented. It was also noted that the concerns of administrators, such as cost and preservation, are often invisible to users. Xie (2008) followed a similar method and uncovered similar results. Usability was again identified as the most important factor to users. In particular, the ease of use and learnability of user interfaces was identified as the single most important factor, although system performance issues (such as response time) were identified as a close second. Another important criterion to the subjects was the presence of feedback mechanisms that allowed users to communicate with administrators in order to point out problems or suggest additional features.

Somerville and Brar (2009) documented the development of an academic digital library project that incorporated users in both the design and evaluation phases. Students under faculty supervision observed the information needs of their peers and created a prototype for a digital library. They then developed user-centered evaluation criteria and used that feedback to perfect the prototype and implement a stable release. These mechanisms were left in place to further continuous improvement of functionality. The participating librarians observed that while they, like most information professionals, enjoyed searching for information, the students valued finding, sharing, and using information. By incorporating their user population from the earliest stages of the project they were able to control for their own biases and create a service that better met the needs of their users. Theng et al. (2000) adopted a similar method of user-centered design by incorporating a primary school class in the development of a digital library for children. A focus group was conducted to determine what features the children would most like to see. Three prototypes were constructed and usability tests were conducted with the class, and then a final focus group was used to determine which prototype the children liked best. The results defied expectation: While children did want a system that was fun, they most valued the same usability features that adults prefer. In particular, a prototype with a virtual bookshelf was preferred to a game-based model as that interface made it easier to find content that they wanted.

Tsakonas and Papatheodorou (2008) was certainly not the only study that revealed how matters of appearance can impact the usability of a system. Zimmer-
man and Paschal (2009) engaged 18 college-aged subjects to evaluate the usability of two digital libraries. Among other findings, participants noted that the font was small enough to make legibility a problem, and the senior researcher added that he had observed many similar comments in usability testing over the last five years. Cunningham and Bennett (2009) demonstrated how even a slight change such as larger font and simplified terminology can vastly increase usability with their redesign of a Greenstone digital library for seniors who wish to archive their personal documents.

### 2.2 Usefulness

Khaltarkhuu and Maeda (2008) explored issues in usefulness through the development of a digital text library in the traditional Mongolian language. This collection has the virtue of preserving a historic language as written Mongolian has made use of the Cyrillic alphabet since 1946. It also protects historical items by removing the need for physical handling (some of the items in the collection are over 1,000 years old) while increasing access to scholars around the world. Thelwall and Vaughan (2004) also examined the usefulness needs of a scholarly user population through an analysis of the Internet Archive. This is a unique digital library that seeks to copy and index as much of the Web as possible and store it indefinitely. This is accomplished by means of Web crawling, a discovery method similar to that used by Google. The researchers estimated that approximately 16% of publicly accessible Web pages are captured in this manner (an impressive amount that consists of many billions of pages). The use of crawlers causes similar coverage issues to the Google search engine, as an HTML link to a page must be placed in a location known to the crawler before it can find the new material. The investigators determined that there is some bias in national origin, but this is inevitable as pages originating from countries that are late adopters of the web will not have as many inlinks as pages from early adopters. No language bias was detected, and the study concluded that coverage is fair enough to provide historians with a rich source of primary documents. It was also noted that the collection is useful for looking up links that no longer point to valid Web addresses.

Kwak and Bae (2009) demonstrated how digital libraries can play a unique role in providing useful content to a user group with special needs. The study made use of online surveys and in-depth interviews to determine the best means of increasing information access for blind users. They identified a greater need in this group than in the general population as they learned that the blind read much
more on average than the non-handicapped. The most useful content type noted was audio files delivered through what the study terms an “ubiquitous” computer network, i.e. one that is not platform dependent. Most participants remarked that mobile phones equipped with sound interaction devices were easier to use than desktop or laptop computers, even those with similar aids to access. The researchers also noted that mobile telephony was a much more useful means of interacting with a digital library than a more traditional telephone service as it allowed users to save the audio files to their phones and transfer them to external computers and storage devices.

Inskip, Butterworth, and MacFarlane (2008) noted that the majority of literature on digital libraries and electronic information seeking focuses on the needs of specialists and investigated the needs of users at a small folk music special library. Participants in the study ranged in skill level from academic users to youth enthusiasts, and their interactions with the digital library system and information needs were determined through focused interviews. The investigators determined that the interface features needed for usability were largely determined by the user’s level of research experience. Academic users were familiar with formulating search queries and made heavy use of search bars, but less experienced users relied on browsing features to navigate the system. The study found that the presence of a digital collection enhanced both usability and usefulness of the library most for the lesser experienced users. This segment of the user population was more likely to be intimidated by finding information in physical formats, especially microfiche, and relied more heavily on the services of librarians. They found that they enjoyed the serendipitous discovery afforded by the browsing features of the navigation interface and felt that the ability to electronically interact with librarians enhanced their access to both the digital and physical collections.

3 Conclusions

In addition to noting the central role of librarians to providing access to obscure information, Inskip et al. (2008) noted that librarians play an even more crucial role to increasing access to digital libraries given the always remote and sometimes voluminous nature of such collections. The investigators further concluded that the role of the digital librarian, much like the role of digital libraries themselves, is still in its infancy. Similar conclusions were drawn by Choi and Rasmussen (2009). Their study analyzed 363 job advertisements for digital librarians placed from 1999 to 2007. Of note was the fact that many positions could not fit neatly
into the definitions of public services or technical services, suggesting that what is emerging is a completely new paradigm for librarianship. Also of note was the fact that managerial qualifications for digital librarian candidates were in very high demand. This suggests that digital librarians are heavily engaged in managing change and transition at the time of this writing, and the new third branch of library services continues to undergo the process of chrysalis and crystallization. 79.31% of positions advertised required an ALA-accredited MLS degree, demonstrating that the required skill-set for digital libraries is shifting from engineering to information services. Further evidence of this transition was found in requirements of technical skills, as only 17.24% of positions required knowledge of HTML and/or Web authoring tools, 13.79% required knowledge of XML or other mark-up languages, and 10.34% required skill in object-oriented programming languages such as Perl, Java, or C++. This trend might change as digital libraries become more organizationally and administratively stable, but in the beginning service commitment trumps technical wizardry in digital librarianship.

Seventeen years after the inauguration of the Digital Libraries Initiative (Borgman, 2003) digital libraries have proven that they will play a central role in providing access to information across the coming Global Information Infrastructure, despite the fact that the nature of this role remains fairly nebulous. Indeed, even defining the term digital library is an exercise somewhat akin to the parable of the three blind men and the elephant. To librarians digital libraries are electronic extensions of service-based institutions that have existed for millennia, to computer scientists they are another of many distributed multimedia information systems, and to end-users they are walled gardens in the wilds of the World Wide Web (Xie, 2006). The information services are no stranger to such ambiguity, however, as the multiplicity of means of meeting information needs are as vast and complex as the concept of information itself. No matter what their precise nature or characteristics, one feature is common to all of the systems that serve users in the role of digital libraries; the provision of access to information through networked systems (Tsakonas & Papatheodorou, 2006).

There can be no provision of library services without librarians, however. It strains the imagination of all but the most ardent of technophiles to picture any automated system that can provide the same level of information access as a trained and skilled librarian. As the Web continues to provide more fuel to the exponentially growing bonfire of the information revolution, the electronic information needs of users will continue to grow in a similar manner. In order to meet their ethical obligation to “provide the highest level of service to all library users” (ALA, 2008), librarians must embrace the change and the opportunities for growth that
change brings with it. The digital divide between the information rich and the information poor might seem at present a gaping chasm, but the growth in rates of access suggests that the coming global digital library might very well represent the last best hope of providing true access for all. Borgman (2003) noted that technological means such as the global digital library in and of themselves will never eliminate inequity in information access. I would add that, with the right training and total commitment, digital librarians just might.

4 References


