What Is Usability in the Context of the Digital Library and How Can It Be Measured?

This paper reviews how usability has been defined in the context of the digital library, what methods have been applied and their applicability, and proposes an evaluation model and a suite of instruments for evaluating usability for academic digital libraries. The model examines effectiveness, efficiency, satisfaction, and learnability. It is found that there exists an interlocking relationship among effectiveness, efficiency, and satisfaction. It also examines how learnability interacts with these three attributes.

D igital library development, since its inception in the 1990s, has made significant progress thus far. Although there is still a long way to go before reaching their full potential, digital libraries are maturing (Fox 2002; Marcum 2002). However, the evaluation of digital libraries has not kept pace. As Saracevic (2000) has outlined, fundamental concepts remain to be clarified, such as What is a digital library? What is there to evaluate? What are the criteria? How to apply them in evaluation? Why evaluate digital libraries in the first place? Borgman (2002) also stated that the digital libraries research community needs large test beds, including collections and testing mechanisms, as a means to evaluate new concepts. There is also a need of benchmarks for comparison between systems and services.

This research is to develop and evaluate methods and instruments for assessing the usability of digital libraries. Compared to other areas in digital library research, as Theng, Mohd-Nasir, and Thimbleby (2000a, 238) point out, “Little work is being done to understand the purpose and usability of digital libraries.” Borgman et al. (2000, 229) also state, “Relatively little work has been done on evaluating the usability of digital libraries in any context.” The same observations are also made by Blandford, Stelmaszewska, and Bryan-Kinns (2003) as well as Brogan (2003). Blandford and Buchanan (2002b) call for a need for further work on methods for analyzing usability, including an understanding of how to balance rigor, appropriateness of techniques, and practical limitations.

This study contributes to the literature the understanding of usability, reviews what methods have been applied and their applicability, and proposes a suite of methods for evaluating usability for academic digital libraries.

Definition of Digital Library

There are many different views in the literature on what digital libraries are. This paper does not intend to provide a comprehensive collection on the definitions of digital libraries, but rather representative ones.

Lesk (1997, 1) views digital libraries as “organized collections of digital information.” Arms (2000, 2) views digital libraries as “managed collection of information, with associated services, where the information is stored in digital formats and accessible over a network.”

The Digital Library Federation (1999) representing the practical community, defines digital library as follows:

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.

Francisco-Revilla et al. (2001) report digital libraries are increasingly being defined as ones that collect pointers to Web-based resources rather than hold the resources themselves. A library’s Web site is an example of this definition. Greenstein (2000) shares this view and says that the digital library is known less for the extent and nature of the collections it owns than for the networked information space it defines through its online services. Paepcke et al. (1996) also state that a digital library provides a single point of access to a wide range of autonomously distributed sources.

In addition, digital libraries may be seen as new forms of information institutions, multimedia information retrieval systems, or information systems that support the creation, use, and searching of digital content (Borgman 2002). Digital libraries also represent a new infrastructure and environment that has been created by the integration and use of computing, communications, and digital content on a global scale destined to become an essential part of the information infrastructure in the twenty-first century (DELOS 2004).

In summary, digital libraries:

- are an organized and managed collection of digital information;
- are accessible over a network; and
- may include service.

Judy Jeng (judyjeng@scils.rutgers.edu) is a Ph.D. candidate at the School of Communication, Information, and Library Studies, Rutgers, The State University of New Jersey.
Dimensions of Usability

Usability is a multidimensional construct that can be examined from various perspectives. The term usability has been used broadly and means different things to different people. Some relate usability to ease of use or user-friendliness and consider from an interface effectiveness point-of-view. This view makes sense, as usability has a theoretical base on human-computer interaction. Many studies on usability focus on interface design. Kim (2002, 26), for instance, points out that "the difference between interface effectiveness and usability is not clear."

Usability can also be related to usefulness and usableness. Gluck (1997), for instance, made this assessment. Usableness refers to such functions as "Can I turn it on?" Can I invoke that function?" Usefulness refers to such functions as "Did it really help me?" Was it worth the effort?" Landauer (1995) distinguishes usability (ease of operation) from usefulness (serving an intended purpose), commenting that the two are hard to separate in the context of evaluation.

Usability has several attributes. The International Standards Organization (1994, 10) defines usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use." Nielsen (1993) points out that usability has five attributes: learnability, efficiency, memorability, low error rate or easy error recovery, and satisfaction. Brinck, Gergle, and Wood (2002) share a similar perspective that usability is functionally correct, efficient to use, easy to learn and remember, error tolerant, and subjectively pleasing. In addition, Booth (1989) outlines that usability has four factors: usefulness, effectiveness (ease of use), learnability, and attitude (likeability). Hix and Hartson (1993) classify usability into initial performance, long-term performance, learnability, retainability, advanced feature usage, first impression, and long-term user satisfaction. Hix and Hartson are unique in that they take one step further to differentiate performance and satisfaction into initial and long-term measures. The definitions given by ISO and Nielsen are most widely cited.

Usability can also be grouped into two large categories: inherent usability (Kurosu and Kashimura 1995) and apparent usability (Kurosu and Kashimura 1995; Tractinsky 1997). Inherent usability is mainly related to the functional or dynamic part of interface usability. It includes those attributes that focus on how to make the product easy to understand, easy to learn, efficient to use, less erroneous, and pleasurable. On the other hand, apparent usability is more related to the visual impression of the interface. At times, inherent usability and apparent usability may be contradictory (Fu 1999). For example, in Web page design, graphics enhance apparent usability but slow down the system.

Usability has user focus. Dumas and Redish (1993, 4) define usability as "people who use the product can do so quickly and easily to accomplish their task." Clairmont, Dickstein, and Mills (1999) make the similar statement that "Usability is the degree to which a user can successfully learn and use a product to achieve a goal."

Usability is different from functionality. Dumas and Redish (1993) use the videocassette recorder (VCR) as an example to illustrate the difference between the two: VCRs may have high functionality (the feature works as it was designed to work) but they have low usability (people cannot use them quickly and easily to accomplish their task). Usability has several aspects, including interface design, functional design, data and metadata, and computer systems and networks (Arms 2000). Usability is a property of the total system. All the components must work together smoothly to create an effective and convenient digital library.

Usability can be tackled from various directions. Blandford and Buchanan (2002a) suggest that usability is technical, cognitive, social, and design-oriented, and it is important to bring these different perspectives together, to share views, experiences, and insights. Indeed, digital library development involves interplay between people, organization, and technology. The usability issue should look at the system as a whole.

In addition to those views, usability can also be examined from the perspectives of graphic design, navigation, and content (Spool et al. 1999). Turner (2002) categorizes usability into navigation, page design, content, accessibility, media use, interactivity, and consistency.

Figure 1 compares various perspectives on the attributes of usability.

Evaluation of Usability

There are a number of ways to evaluate usability. The techniques include formal usability testing; usability inspection; card sort; category membership expectation; focus groups; questionnaires; think-aloud; analysis of site usage logs; cognitive walkthrough; heuristic evaluation; claims analysis; concept-based analysis of surface and structural misfits (CASSM); and paper prototyping (Askin 1998; Blandford et al. 2004; Campbell 2001; Kantner and Rosenbaum 1997; Keith et al. 2003; Nielsen 1993; Redish 1993; Tractinsky 1997).
The areas of usability testing for digital libraries have covered breadth of coverage, navigation, functionality, utility, interface, metadata appropriateness, and awareness of library resources.

The National Taiwan University Library used questionnaires to survey 1,784 users on usability (Lan 2001). They found the site’s usability problems are mainly in the areas of information architecture and in the browsing and searching mechanism. The study of CUNY+ (Oulanov and Pajarillo 2002) also employed a questionnaire as the primary method of usability assessment. The authors conducted a two-phase study to compare usability of text-based and Web-based CUNY Web sites. The criteria used were affect, efficiency, control, helpfulness, and adaptability.

Adams and Blandford (2002) reported on their study of accessibility on a large London-based hospital. They conducted focus groups and in-depth interviews with seventy-three hospital clinicians. Fifty percent of the participants were nurses, while the other fifty percent were senior and junior doctors, consultants, surgeons, managers, and IT department members. The study focused on two themes: (1) the perceived effectiveness of traditional and digital libraries as clinical resources; and (2) the impact of clinician status on control over and access to information. Participants responded that digital library technology provides remote access to materials, but the system’s usability is poor and it is time-consuming to access information.

Theng, Mohd-Nasir, and Thimbleby (2000a) utilized questionnaires and heuristic evaluation to measure usability of the ACM Digital Library, the Networked Computer Science Technical Reference Library, and the New Zealand Digital Library. This study helps to understand the purpose of digital libraries.

Sumner and Dawe (2001) studied usability of the Digital Library for Earth System Education (DLESE) focusing on its role in the process of educational resource reuse. One finding is that the design of the search results page is critical for supporting resource comprehension. Also, the library’s metadata plays a central role in documenting the resource enough to support comprehension and modification processes.

Sumner et al. (2003) again used DLESE to study usability in addition to National Science Digital Library (NSDL). The purpose of this study was to identify educators’ expectations and requirements for the design of educational digital collections for classroom use. A series of five focus groups was conducted with a total of thirty-six teachers and two librarians to review eighteen Web sites. The participants indicated that content quality, advertising, bias, and design were important factors influencing their perceptions.

Hartson, Shivakumar, and Pérez-Quiñones (2004) applied the usability inspection method to evaluate the design and functionality of Networked Computer Science Technical Reference Library (NCSTRL). They found NCSTRL’s design was apparently functionally oriented rather than an approach based on user task threads. Another finding of the usability inspection was about terminology used in NCSTRL. There was jargon and the use of terms was designer-centered rather than user-centered.

The evaluation of the University of Illinois DeLiver service applied a mix of methods, including transaction log analysis, surveys, interviews, focus groups, and formal usability testing to measure accessibility (Bishop...
The proposed evaluation model applies the definition of ISO 9241-11 (International Standards Organization, 1994) that examines effectiveness, efficiency, and satisfaction. In addition, the model includes learnability (see figure 3). The ISO definition defines usability as “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.” (10) The ISO definition, however, does not explicitly specify operational criteria on what to evaluate.

In the proposed model, effectiveness is evaluated by whether the system as a whole can provide information and functionality effectively and will be measured by how many answers are correct. Efficiency is likewise evaluated by the system’s ability to retrieve information efficiently and will be measured by how much time it takes to complete tasks. Satisfaction will look into the areas of ease of use, organization of information, clear labeling, visual appearance, contents, and error corrections and will be measured by Likert scales and questionnaires. Ease of use evaluates a user’s perceptions about the ease of use of the system. Organization of information evaluates whether the system’s structure, layout, and organization meets the user’s satisfaction. Labeling examines from the user’s perception whether the system provides clear labeling and if terminology used is easy to understand. Visual appearance evaluates the site’s design to see if it is visually attractive. Content evaluates the authority and accuracy of information provided. Error tests whether users recover from mistakes easily and if they make mistakes easily due to system’s design. Learnability measures learning effort. The learning effort takes into consideration how soon the subject begins to know how to perform tasks and how many tasks are completed correctly.

Figure 3 is a diagram illustrating this evaluation model. It is suspected that there exists an interlocking relationship among effectiveness, efficiency, and satisfaction. In addition, it will be interesting to examine how learnability interacts with effectiveness, efficiency, and satisfaction.

**Usability Evaluation Instruments**

A set of instruments are designed based on the evaluation model. University and college library Web sites are selected as an example to test the model and instruments for the purpose of this paper. The instruments include a pretest questionnaire (see appendix A), a list of tasks (see appendix B), and a post-test questionnaire (see appendix C).

The pretest questionnaire collects demographic data, including gender, age, status (undergraduate, master’s, or doctoral student), major, years at the institution, original nationality if coming from a foreign country, and familiarity with the site. There have been studies on how gender, age, and cultural differences affect how people interact with online information (Collins and Auguiñaga 2001; Duncker 2002; Vohringer-Kuhnt 2003). A university or college library Web site serves a diverse student body, including international students and students in a wide range of ages. It is interesting to examine how those demographic factors influence usability assessment.
The list of tasks includes nine questions that are representative of typical uses of a library’s Web site. Three of those questions are to locate known items, including author, title, and e-book searching. Four are to use databases to find articles in electronic journals. Two are to locate information, such as eligibility for ILL services and how to set up remote access.

The subjects are asked to rank satisfaction with the system after each task and to write down comments. In addition, there is a post-test questionnaire that specifically

<table>
<thead>
<tr>
<th>Site</th>
<th>Methods</th>
<th>Subjects</th>
<th>Areas</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACM, IEEE-CS, NCSTRL, NDLTD</td>
<td>formal usability test, questionnaire</td>
<td>48 students (38 graduate, 10 undergraduate)</td>
<td>interface</td>
<td>Kengeri et al. (1999)</td>
</tr>
<tr>
<td>ACMDL, NCSTRL, NZDL</td>
<td>questionnaire, heuristic evaluation</td>
<td>45 undergraduate</td>
<td>design and structure</td>
<td>Theng et al. (2000a, 2000b)</td>
</tr>
<tr>
<td>Alexandria</td>
<td>questionnaire, formal usability test</td>
<td>23 students</td>
<td>interface</td>
<td>Thomas (1998)</td>
</tr>
<tr>
<td>CUNY+</td>
<td>questionnaire</td>
<td>10 students</td>
<td>interface</td>
<td>Oulanov &amp; Pajarillo (2002)</td>
</tr>
<tr>
<td>DeLiver</td>
<td>transaction log survey, interview, focus groups, formal usability test</td>
<td>1900 graduate, 420 faculty</td>
<td>accessibility</td>
<td>Neumann &amp; Bishop (1998), Bishop (2001)</td>
</tr>
<tr>
<td>DLESE, NSDL</td>
<td>focus groups</td>
<td>36 teachers, 2 librarians</td>
<td>design</td>
<td>Sumner et al. (2003)</td>
</tr>
<tr>
<td>Instructional Architect</td>
<td>formal usability test, focus group</td>
<td>26 teachers</td>
<td>interface, content</td>
<td>Dorward et al. (2002)</td>
</tr>
<tr>
<td>London Hospital</td>
<td>focus groups, interviews</td>
<td>73 clinicians</td>
<td>accessibility</td>
<td>Adams &amp; Blandford (2002)</td>
</tr>
<tr>
<td>MARIAN (Virginia Tech)</td>
<td>formal usability test, log analysis, questionnaire</td>
<td>students, faculty, staff</td>
<td>interface</td>
<td>France et al. (1999)</td>
</tr>
<tr>
<td>MIT</td>
<td>formal usability test</td>
<td>29 (faculty, graduate, undergraduate, staff)</td>
<td>site design</td>
<td>Hennig (1999)</td>
</tr>
<tr>
<td>National Taiwan U.</td>
<td>questionnaire</td>
<td>1784 faculty and students</td>
<td>information architecture, browsing &amp; searching mechanism, layout and display</td>
<td>Lan (2001)</td>
</tr>
<tr>
<td>NCSTRL</td>
<td>usability inspection</td>
<td>3 usability experts</td>
<td>design, interface, functionality</td>
<td>Hartson et al. (2004)</td>
</tr>
<tr>
<td>SABIO</td>
<td>formal usability test, heuristic evaluation, design walk-through, card sorting</td>
<td>students</td>
<td>design</td>
<td>Dickstein &amp; Mills (2000)</td>
</tr>
<tr>
<td>U. of Illinois at Chicago</td>
<td>formal usability test</td>
<td>12 students</td>
<td>navigation</td>
<td>Augustine &amp; Greene (2002)</td>
</tr>
<tr>
<td>U. of South Florida</td>
<td>formal usability test</td>
<td>26 undergraduate</td>
<td>interface</td>
<td>Allen (2002)</td>
</tr>
<tr>
<td>U. of the Pacific</td>
<td>formal usability test</td>
<td>134 students</td>
<td>awareness of library resources</td>
<td>Krueger et al. (2004)</td>
</tr>
<tr>
<td>Washington State U.</td>
<td>formal usability test, questionnaire</td>
<td>12 students</td>
<td>navigation</td>
<td>Chisman et al. (1999)</td>
</tr>
</tbody>
</table>

**Figure 2. Methods of usability evaluation**
examines satisfaction in the areas of ease of use, organization of information, clear labeling, visual appearance, contents, and error corrections.

Testing of the Model and Instruments

The earlier version of the model and instruments was tested using three students at the Rutgers University Libraries Web site. Revisions were made after the pilot study. The current version of the model and instruments are tested at the Rutgers University Libraries Web site (www.libraries.rutgers.edu) and the Queens College Library Web site (http://qcpages.qc.edu/Library). It is hoped that the model and instruments can be generalized for use in academic digital libraries.

The study employs a number of techniques, including formal usability testing, questionnaire, interview, think aloud, and log analysis. The evaluation model and instruments in this study consider both the quantifying elements of performance (time, accuracy rate, steps to complete tasks) as well as subjective criteria (satisfaction). Satisfaction is further examined in the areas of ease of use, organization of information, labeling, visual appearance, content, and error correction. The evaluation approach is empirical.

Results

While the primary interest of this study is to devise an evaluation model and a suite of instruments for evaluating usability of academic digital libraries, the data collected in the study are used to explore the following usability issues.

Literature review has indicated that there is a need of usability testing benchmarks for comparison. For example, Theng, Mohd-Nasir, and Thimbleby (2000b) report that they had to make the assumption that if an area scores 75 percent and above for accuracy it implies that the area is well implemented. The usability testing at MIT Libraries also report that subjects had 75 percent success rate (Hennig 1999). But, they wondered, is 75 percent high or low? The results of the usability testing of this study are forthcoming in the author’s doctoral dissertation and will be contributed to the literature as a benchmark.

In addition, this research examines the issues of user lostness and navigation disorientation. The user lostness issue has been reported by several studies, including Blandford, Stelmaszewski, and Bryan-Kinns (2001), Buttenfield (1999), Gullikson et al. (1999), Kengeri et al. (1999), and Spool et al. (1999) as well as by Theng, Mohd-Nasir, and Thimbleby (2000a). Indeed, navigation disorientation is among the biggest frustrations for Web users (Brinck, Gergle, and Wood 2002). This situation is common particularly with the increasing provision of digital library portals that provide links to various libraries from one Web site.

This research also examines if there exists an interrelated relationship among effectiveness, efficiency, and satisfaction. The results indicate this relationship. The results are reported in Jeng (2004) and will also be available in the author’s doctoral dissertation. Although there is an interlocking relationship among these three criteria, each has its own emphasis and should be measured separately.

Contribution

This paper contributes to the literature an evaluation model and a suite of instruments for evaluating usability of academic digital libraries. It calls attention to the potential usability differences due to age and culture, the user lostness and navigation disorientation issues,
and the need for benchmarks. It discusses usability in the context of digital libraries and examines how it has been evaluated. This study will continue as doctoral dissertation research, and the results will be shared with academic and professional communities.

Editor’s note: Ms. Jeng’s article is the winner of the 2004 LITA/Endeavor Student Writing Award.

References


Campbell, Nicole. 2001. Usability assessment of library-related Web sites: Methods and case studies. Chicago: ALA.


Appendix A. Pretest Questionnaire

Thank you very much for agreeing to participate in this experiment. All of your personal data that we collect will be entirely confidential, viewed only by the experimenter, and shared only as part of group results. But first, we would like to gather a bit of background information about you, so that we will be better able to interpret your use of and reactions to the system.

Participant # ______
Date: _________________
Gender: ___ Male ___ Female
Age: _____
What is your current status:
_____ Undergraduate _____ Master’s Student
_____ Doctoral Student _____ Faculty
Major/Department: _______________________________
How many years have you been at Rutgers or Queens? _______
If you are from foreign country, how long have you been in the U.S.? ________ years
Your original nationality: _________________________
Ethnic group: ___ White ___ African American ___ Asian
___ Hispanic ___ Native American ___ Other: _________
How often do you use the Library’s Web site:
_____ Never used it
_____ Once or twice a semester
_____ Once or twice a month
_____ Once or twice a week
_____ Daily

Appendix B. Usability Testing Questions

The goal of this test is to evaluate the usability of the library’s Web site. I will ask you a series of questions and would like you to think out loud while you look for the answer. Some questions are easy and some are more difficult. Do not worry if you can’t find the answer every time. Please remember that we are testing the effectiveness of the site design and this is not a test of you. The whole test should take less than an hour. I thank you.

1. Does the library have a copy of Gone with the Wind, book format, by Margaret Mitchell?

Please rank from 1 to 5 regarding the ease of use of the system, 1 being the easiest and 5 being the most difficult.

1  2  3  4  5
Easy to use               Difficult to use
Your comment: ________________________________

2. Does the library currently subscribe to paper copy of Advertising Age?

1  2  3  4  5
Easy to use               Difficult to use
Your comment: ________________________________

3. Use a database to find an article about nursing homes and mental illness.

1  2  3  4  5
Easy to use               Difficult to use
Your comment: ________________________________

4. Find a journal article on gospel music.

1  2  3  4  5
Easy to use               Difficult to use
Your comment: ________________________________

5. I am interested in investing in what are referred to as “callable securities.” Please find a recent article about them.

1  2  3  4  5
Easy to use               Difficult to use
Your comment: ________________________________

6. Find an encyclopedia article about French wine.

1  2  3  4  5
Easy to use               Difficult to use
Your comment: ________________________________
Appendix C. Post-Test Questionnaire

Thanks again for participating in this experiment. This questionnaire gives you an opportunity to tell us your reactions to the system you used. Please circle a number on the scale to indicate your reactions. Please write comments to elaborate on your answers. I will go over your answers with you to make sure that I understand all of your responses. Thank you.

1. Please rate the ease of use of the Web site.
   1  2  3  4  5
   Easy                 Difficult
Your comment: ___________________________________
__________________________________________________

2. What do you think about the organization of information on the site?
   1  2  3  4  5
   Clear                  Unclear
Your comment: ___________________________________
__________________________________________________

3. What do you think about the terminology used in the site? Are categories clearly labeled?
   1  2  3  4  5
   Clear                  Unclear
Your comment: ___________________________________
__________________________________________________

4. Is the site visually attractive?
   1  2  3  4  5
   Attractive               Unattractive
Your comment: ___________________________________
__________________________________________________

5. What is the best feature(s) of the site? _______________
   __________________________________________________

6. What is the worst feature(s) of the site? _______________
   __________________________________________________

7. What new content or features that you would like to see on the site? _______________
   __________________________________________________

8. Can you recover from mistakes easily?
   1  2  3  4  5
   Easy                 Difficult
Your comment: ___________________________________
__________________________________________________

9. Your overall reaction to the system:
   1  2  3  4  5
   Satisfied                 Unsatisfied
Your comment: ___________________________________
__________________________________________________

10. Do you feel lost while using the site?
    ____ Yes  _____ No
    Your comment: ___________________________________
__________________________________________________

11. Is the site easy to navigate?
    ____ Yes  _____ No
    Your comment: ___________________________________
__________________________________________________

12. When you click a button on the Web page, do you expect that the click will lead you to correct answer?
    ____ Yes  _____ No
    Your comment: ___________________________________
__________________________________________________

13. Do you have any other comments about the Web site?
    __________________________________________________
    __________________________________________________
    __________________________________________________
    __________________________________________________