Digital Library Evaluation Criteria: What do Users Want?*

디지털 도서관 평가기준: 이용자들이 원하는 것은 무엇인가?

Iris Xie**

Soohyung Joo***

Krystyna K. Matusiak****

ABSTRACT

Criteria for evaluating digital libraries have been suggested in prior studies, but limited research has been done to understand users' perceptions of evaluation criteria. This study investigates users' opinions of the importance of digital library evaluation criteria. Thirty user participants, including 10 faculty members and 20 students, were recruited from five universities across the United States. They were asked to rate the importance of evaluation criteria in eight dimensions (e.g. collections, information organization, context, etc.). The results demonstrate that users care about evaluation criteria related to the use and quality of collection and services rather than the operation of digital libraries. The findings of this study are relevant to the development of user-centered digital libraries and associated evaluation frameworks through the incorporation of unique users' needs and preferences.

초 록

기존의 여러 연구들에서 디지털 도서관 평가를 위한 기준들이 제시되어 왔으나, 대부분의 연구에서 평가기준에 대한 이용자 관점을 이해하려는 노력이 부족하였다. 본 연구는 디지털 도서관 평가 기준들에 대한 이용자들의 의견을 미국 내 5개 대학교에서 10명의 교수 이용자와 20명의 학생 이용자에게 설문을 통해 직접 조사하였다. 설문 참여자들은 본 연구진에 의해 제시된 8개 영역 (장서, 정보조직, 맥락 등) 내 평가 기준들의 중요성에 대해 7점 척도로 응답하였다. 설문의 결과는 이용자들이 장서의 이용과 품질, 서비스와 관련된 항목들을 도서관 운영항목에 비해 상대적으로 더 중요하게 생각하고 있었음을 보여주었다. 본 연구의 결과는 이용자 중심의 디지털 도서관 개발과 관련하여 이용자의 요구와 선호를 반영하는 도서관 평가체계 구축에 도움이 될 것이다.

Keywords: digital libraries, digital library evaluation, evaluation criteria, user perspectives, evaluation dimensions
디지털 도서관, 디지털 도서관 평가, 평가 기준, 이용자 관점, 평가 영역

1. Introduction and Literature Review

Digital libraries (DLs) have emerged as one of the essential scholarly information systems in support of research and teaching. Many libraries have digitized their collections, such as pictures, books, or audios, to make them available on the web. Digital libraries are relatively new phenomena and, like many new and emergent information systems, face challenges of acceptance, utilization, and evaluation. The concept of a digital library has been perceived in different ways by different groups of people. To the community of librarians and LIS researchers, a digital library is an extension and augmentation of library services combined with a remote access to digitized resources. To computer scientists, a digital library is a distributed information system or networked multimedia information system (Fox et al. 1995). In this study, our focus is on the users' perspective. A digital library is defined as the collection of digitized or digitally born items that are stored, managed, serviced, and preserved by digital library professionals.

The exponential growth of DLs has created a need for the evaluation of these emergent information systems. Digital library evaluation has become a critical issue for both professionals and researchers. DLs have their own unique characteristics and features compared to traditional library services. Also, as an information retrieval system, digital libraries are fairly different from other types of online information retrieval systems (e.g., search engines,

online databases, OPACs). DLs have well-organized metadata, browsing categories, as well as digitized items in different formats. Therefore, previous evaluation frameworks used in traditional libraries or other information retrieval systems proved to be insufficient in assessing different aspects of DLs. The evaluation of DLs is conceptually complex and pragmatically challenging (Saracevic and Covi 2000). Borgman et al. (2000) also pointed out technical complexity, variety of content, and the lack of evaluation methods posed the key challenges of digital library evaluation.

Researchers have exerted efforts to develop new frameworks and methods of digital library evaluation. A number of evaluation criteria were suggested covering different dimensions of digital libraries. The early DL projects, funded by the National Science Foundation (NSF) as part of Digital Libraries Initiatives I and II, laid groundwork in evaluation research by producing DL prototypes and frameworks (Borgman et al. 2000; Buttenfield 1999; Hill et al. 2000; Van House et al. 1996). In particular, Hill et al. (1997) identified several criteria for digital library evaluation, such as ease of use, overall appeal, usefulness and overall performance. Saracevic (2004) identified six classes of criteria: content, technology, interface, process/service, user, and context. This evaluation framework covers multiple aspects of digital libraries comprehensively, and is one of the first attempts to assess the context aspect in digital libraries. Xie's (2006/2008) evaluation framework shifted a focus to the users and posited five types of criteria: usability, collection quality,

service quality, system performance efficiency, and user feedback solicitation. In particular, when developing the evaluation framework, she analyzed users' perceptions based on diaries, questionnaires, and interviews. Zhang (2010) validated the evaluation criteria for digital libraries. Based on Saracevic's (2004) framework, she investigated the importance of evaluation criteria using the empirical survey data from heterogeneous stakeholders. Noh (2010) identified multiple dimensions and corresponding evaluation indices for electronic resources.

In Europe, DELOS Network of Excellence has conducted a series of projects regarding the evaluation of digital libraries. DELOS is a comprehensive and large scale DL project, which represents joint activities aimed at integrating and coordinating the ongoing research efforts of the major European teams working in the digital library area. Candela et al. (2007) established DELOS Manifesto which presents a three-tier DL framework incorporating six core components such as content, functionality, quality, policy, architecture, and user. Fuhr et al. (2001) proposed an evaluation scheme for digital libraries which covers four dimensions including data/collection, system/technology, users, and usage. Based on the examination of the interactions amongst digital library components, Tsakonas et al. (2004) proposed the major evaluation foci, such as usability, usefulness, and system performance. Fuhr et al. (2007) developed a DL evaluation framework based on a DELOS model and conducted a large-scale survey of DL evaluation activities.

Among different aspects of digital libraries, us-ability has been one of the major concerns in evaluation. Usability consists of multiple attributes from various perspectives such as learnability, efficiency, effectiveness, memorability, errors, and satisfaction (Nielson 1993). Several researchers tried to suggest a usability evaluation model tailored to digital libraries. For example, Jeng (2005) suggested an evaluation framework on usability of academic DLs focusing on four attributes: effectiveness, efficiency, satisfaction, and learnability. Ward and Hiller (2005) suggested usability evaluation criteria specific to library services, such as completion of the task, time and effort, and reaction to the product or service. Joo and Lee (2011) developed an instrument tool to measure the usability of digital libraries. They further tested the validity and reliability of the tool. Matusiak (2012) examined the relationship between usability and usefulness, and found that user perceptions of usefulness and usability, especially perceived ease of use, play an important role in user intentions to adopt and use digital collections.

Researchers strived to identify a set of evaluation criteria for digital libraries. However, there has been relatively less effort devoted to the investigation of users' perceptions in selecting evaluation criteria. Ultimately, digital libraries are developed to provide information and services to users, and users' opinions should be considered in the evaluation of digital libraries. It is important that all efforts in the evaluation of digital libraries should be rooted in users' information needs and characteristics as well as contexts involving the users of those libraries (Marchionini et al. 1998).

This study is one of the few attempts to survey users' perceptions of evaluation criteria for digital libraries. The investigation of user perceptions is a fundamental step in devising an evaluation framework that focuses on user needs and characteristics. In this study, the authors suggested a wide range of evaluation criteria in eight dimensions of digital libraries based on the document analysis. For the suggested evaluation criteria, this study examines to what extent users perceive the importance of each criterion in the evaluation of digital libraries.

2. Research Problem and Research Question

Thus far, digital library evaluation criteria have been suggested mainly by librarians or researchers. To design a user-centered digital library, the evaluation needs to reflect users' perspective in its evaluation criteria. This study is one of a few studies that investigated users' perceptions of evaluation criteria for digital libraries.

This study intends to examine the following research question:

What are users' perceptions of the importance of digital library evaluation criteria?

2.1 METHODOLOGY

Two-round surveys were conducted to identify the importance of evaluation criteria and appropriate-

ness of measures from different stakeholders of digital libraries including scholars, digital librarians, and users. This paper focuses on the identification of the importance of evaluation criteria from users' perspectives.

The authors partnered with five academic libraries across the United States to collect data. Subjects of this study were recruited from these partner libraries: (1) University of Denver, (2) University of Florida, (3) University of Nevada Las Vegas, (4) Drake University, and (5) University of Wisconsin-Milwaukee. Each institution recruited six digital library users to participate in the study. The study employed purposeful sampling strategy. The sample included academic users with prior experience interacting with digital collections. To ensure the maximum variation sampling, participants were recruited from different groups of academic users with different gender and different majors, such as Linguistics, English, Psychology, and Computer Science. The user subjects included 10 faculty, 12 graduate students, and 8 undergraduate students. A $30 gift card was given to each subject as an incentive for his/her participation of the study. Table 1 presents the demographic data of the subjects.

A comprehensive survey was administered to investigate users' perceptions of the importance of evaluation criteria in digital library evaluation. To suggest an initial set of evaluation criteria, a comprehensive document analysis was conducted. Using keywords of different combinations of "digital library", "evaluation", "criteria", and other terms, relevant research papers were collected through Google

〈Table 1〉 Demographic data of subjects

	Category	Frequency	%
Group	Undergraduate	8	26.7%
	Graduate	12	40%
	Faculty	10	33.3%
Gender	Male	12	40%
	Female	18	60%
Student Major	Linguistics, English, Elementary and Special Education, History, Film, Curriculum and Instruction, Geographic Information Sciences, Finance, Geography, Environmental Science, Mental Health Counseling, Art, etc.		
Faculty Major	Counselor Education, English, History, Geography, Computer Science, Law, Psychology, Journalism, Information Science		

Scholar and EBSCOhost databases. In addition, five websites related to digital libraries were also analyzed, such as DigiQUAL and DELOS, to identify evaluation criteria.

Digital library evaluation criteria were extracted from the retrieved pool of documents. Based on the document analysis, ten essential dimensions of digital libraries have been identified, including collection, information organization, interface design, system and technology, effects on users, services, preservation, administration, user engagement, and context. The administration and preservation dimensions were excluded from the user survey because users don't have enough knowledge of these dimensions. For each dimension, the authors proposed a set of evaluation criteria. To help subjects understand the meaning of evaluation criteria, the definition associated with each criterion was provided in the survey. Subjects were instructed to rate the importance of evaluation criteria using a seven point scale. Since the subjects were selected from different locations with different digital library uses, the subjects were not asked to evaluate a specific digital library, but rather based their survey responses on their past interactions with digital library systems.

Descriptive statistics was used, including mean and standard deviation, to investigate the importance of evaluation criteria. Based on average ratings, the authors ranked the evaluation criteria from the most important to the least for each dimension.

2.2 RESULTS

The results section is organized in eight dimensions of digital libraries: collection, information organization, interface design, system and technology, effects on users, services, user engagement, and context. In the dimension of collections, quality related evaluation criteria are the ones that users considered the most important. "Authority (6.53)", "item quality (6.27)", and "digitization standards (6.20)" turned

out to be the top three evaluation criteria. Following these criteria, "cost (6.10)", "format compatibility (6.10)" and "contextual information (6.10)" were ranked fourth. In contrast, "size (5.57)", "diversity (5.77)" and "completeness (5.77)" were considered to be the least important. It seems that users cared more about the quality and less about the comprehensiveness and variety of the collections. Table 2 presents the importance of evaluation criteria in the dimension of collections.

For the dimension of information organization, users perceived metadata as the key. In particular, accuracy and consistency of metadata are the most important criteria in assessing the organization of digital libraries. "Metadata accuracy (6.28)", "consistency (6.24)" and "depth of metadata (6.21)" were ranked 1st, 2nd, and 3rd respectively. "Comprehensiveness (6.10)", "accessibility to metadata (6.07)", and "appropriateness (6.03)" also re-

ceived high scores. On the other hand, users did not consider highly the evaluation criteria that professionals care in developing digital libraries. "Metadata interoperability (5.48)", "controlled vocabulary (5.69)", and "metadata standards (5.86)" were perceived the least important. Table 3 presents the importance of evaluation criteria in the dimension of information organization.

In terms of interface design, the users regarded "browsing function (6.53)" and "search function (6.48)" as the most important criteria in evaluating digital libraries. Searching and browsing are the two main approaches in the information retrieval process. Browsing is a unique feature for digital libraries because of the nature of the digital collections. Users perceived these two criteria as the important criteria in digital library interface design. "Navigation (6.36)" and "reliability (6.28)" were also chosen as important evaluation criteria by the user group. However,

〈Table 2〉 Importance of evaluation criteria in the dimension of collections

Criteria	Mean	STD
Authority (To assess whether information provided by a DL comes from trustworthy sources)	6.53	0.688
Item quality (To assess the quality of a digitized item provided by a DL within its format)	6.27	0.825
Digitization standards (To assess the types and methods of digitization practices conducted, as well as identifying best practices to build a DL)	6.20	0.819
Cost (To assess costs to build collections in a DL)	6.10	0.803
Format compatibility (To assess whether the format of collections in a DL are compatible with a variety of software and systems for different purposes)	6.10	0.774
Contextual information (To assess what type of additional, related information)	6.10	0.799
Audience (To assess who are the main potential users of a DL)	6.07	1.113
Scope/coverage (To assess the range of topics that is covered by a DL)	5.83	0.923
Completeness (To assess whether a DL covers all documents in each topic area)	5.77	1.085
Diversity (To assess whether a DL deals with a variety of issues in relation to a topic of interest)	5.77	1.236
Size (To assess the amount of collection items provided by a DL)	5.57	1.045

〈Table 3〉 Importance of evaluation criteria in the dimension of information organization

Criteria	Mean	STD
Metadata accuracy (To assess how accurately metadata elements are assigned for each item)	6.28	0.854
Consistency (To assess whether metadata is consistent across collections in a DL)	6.24	1.031
Depth of metadata (To assess to which levels of depth metadata elements cover in a DL)	6.21	1.166
Comprehensiveness (To assess whether the organization structure covers all the access points of a DL)	6.10	0.848
Accessibility to metadata (To assess how easily a user can obtain the metadata information of each item)	6.07	1.138
Appropriateness (To assess whether the organizational structure and associated categories adequately organize items in a DL)	6.03	1.036
Metadata standards (To assess whether metadata elements follow predefined standard and guides)	5.86	0.917
Controlled vocabulary (To assess the type and amount of controlled vocabularies used in a DL)	5.69	0.826
Metadata Interoperability (To assess whether metadata elements of a DL are compatible to different DLs)	5.48	1.133

"personalized page (4.17)", "user control (5.14)", and "visual appeal (5.59)" were rated least important in this dimension. In this study, customized features were not deemed as important as assessment criteria. Table 4 presents the importance of evaluation criteria in the dimension of interface design.

As to the dimension of system and technology, effectiveness and reliability of digital libraries are the key evaluation criteria to users. "Response time (6.26)", "retrieval effectiveness (6.25)", and "reliability (6.25)" turned out the most important criteria from the user perspective in DL evaluation. As DLs are considered as one type of the information retrieval systems, the subjects thought that response time and retrieval effectiveness (e.g., precision, recall, etc.) would be important in evaluating the performance of digital library systems. Reliability is a criterion needed to provide stable services to users in digital libraries. "Server performance (5.93)", "fit-to-task (5.93)", and "error rate/ error correction (5.93)" were

tied as the fifth. Less important criteria were "linkage with other digital libraries (5.36)", "integrated search (5.75)", and "flexibility (5.82)". Comparatively speaking, users cared less about the ability to integrate different collections within the digital library environment. Table 5 presents the importance of evaluation criteria in the dimension of system and technology.

In users' ratings of the criteria in the dimension of effects on users, research output and learning effects are essential to users because these are related to their goals in the academic world. They perceived research and learning as the most important aspects of effects on users to be assessed in digital libraries. "Research productivity (5.89)" and "learning effects (5.46)" were chosen as the two most important criteria. Following that, "instructional efficiency (5.32)" and "knowledge change (5.26)" were ranked at third and fourth respectively. On the contrary, evaluation criteria that general users care about are comparatively less important to them. "Information

〈Table 4〉 Importance of evaluation criteria in the dimension of interface design

Criteria	Mean	STD
Browsing function (To assess in what ways and to what extent the interface supports a user's ability to surf related items in a DL)	6.53	0.634
Search function (To assess what types of search functions are offered by a DL and how easy to use them)	6.48	0.693
Navigation (To assess in what ways and to what extent the interface supports a user's exploration in a DL)	6.36	0.920
Reliability (To assess the ability of a DL to perform and maintain its functions under different circumstances)	6.28	0.731
Intuitive operation (To assess how straightforward a DL interface is for a user to understand its operation, and how easily a user can learn to operate the interface)	6.25	1.031
Consistency (To assess whether the design and layout are coherent across a DL interface)	6.14	0.819
Search results presentation (To assess the types of formats/options of search results are presented to users in a DL)	6.10	0.932
Help function (To assess what types of help functions are offered and how effectively they support users in their help-seeking process)	5.93	0.881
Visual appeal (To assess to what extent the interface of a DL is visually attractive to users)	5.59	1.197
User control (To assess to what extent the DL allows users to manipulate its interface)	5.14	1.512
Personalized page (To assess whether a DL offers personalized pages based on user profile)	4.17	1.584

〈Table 5〉 Importance of evaluation criteria in the dimension of system and technology

Criteria	Mean	STD
Response time (To assess how quickly a DL responds to a user's request)	6.26	1.002
Retrieval effectiveness (To assess how effective the search algorithm is in a DL)	6.25	1.023
Reliability (To assess how stable a DL's performance is over time)	6.25	0.823
Connectivity (To assess how stable a DL system can be connected to other information systems)	6.04	0.808
Server performance (To assess the ability of a server to run a DL)	5.93	1.072
Fit-to-task (To assess to what extent a DL is adequate to perform tasks that a user requests)	5.93	0.759
Error rate/ error correction (To assess the degree of errors encountered during the use of content management system and the ability fix the errors)	5.93	0.832
Page loading speed (To assess how quickly a DL presents a user-requested page)	5.86	1.160
Flexibility (To assess whether a DL responds to potential internal or external changes in a timely manner)	5.82	0.818
Integrated search (To assess whether a DL offers an integrated search environment for different collections within a DL)	5.75	0.993
Linkage with other DLs (To assess the identification of and in what ways a DL is linked to other related DLs)	5.36	1.079

literacy/ skill change (5.00)" and "perceptions of digital libraries (5.11)" were regarded relatively less important. Table 6 presents the importance of evaluation criteria in the dimension of effects on users.

In the dimension of services, the subjects again considered reliability and quality of service as important criteria. "Services for users with disabilities (6.43)", "reliability (6.39)" and "service quality (6.36)" were selected as three most important criteria. Interestingly, the subjects thought that the evaluation should reflect types of services tailored to users with disabilities (e.g., blind users, visually impaired users, etc.). The next three criteria (ranked 2nd, 3rd, and 4th) -- "reliability," "service quality" and "user satisfaction" -- are commonly used evaluation criteria for services in other types of information systems. "Usefulness (6.29)", "responsiveness (6.21)", and "timeliness (6.11)" are ranked 5th, 6th, and 7th respectively by showing rating scores above 6. On the other hand, "customized services (4.89)", "types of unique services (5.14)", and "user education (5.36)" were ranked the least important criteria. It

was noted that users thought "customized services" comparatively less important. Apparently, users expected less in regard to special services. Table 7 presents the importance of evaluation criteria in the dimension of services.

In the dimension of user engagement, "user feedback (5.89)", "resource use (5.81)", and "help feature use (5.79)" were the three highly rated evaluation criteria by the user group. "User feedback" is one of the explicit and direct communication channels between users and digital libraries. "Resource use" is one of the fundamental criteria in library assessment, and it is also perceived as an important evaluation criterion in the context of digital library. Since digital libraries represent a new type of IR system, users still need to use help features in order to effectively access digital libraries. On the other hand, "e-commerce support (4.89)" and "user knowledge contribution (5.00)" were perceived less important in evaluation. Table 8 presents the importance of evaluation criteria in the dimension of user engagement.

〈Table 6〉 Importance of evaluation criteria in the dimension of effects on users

Criteria	Mean	STD
Research productivity (To assess in what ways and to what extent a DL affects a user's research outputs)	5.89	1.035
Learning effects (To assess in what ways and to what extent a DL influences a user's learning outcome)	5.46	1.281
Instructional efficiency (To assess in what ways and to what extent a DL enhances user's teaching effectiveness)	5.32	1.144
Knowledge change (To assess in what ways and to what extent a DL influences a user's knowledge structure)	5.26	1.251
Perception of digital libraries (To assess in what ways and to what extent a DL influences users' view of digital libraries)	5.11	1.259
Information literacy/skill change (To asses to what extent DL enhances users' literacy skills)	5.00	1.368

〈Table 7〉 Importance of evaluation criteria in the dimension of services

Criteria	Mean	STD
Services for users w/ disabilities (To assess the types of services offered to users with disabilities)	6.43	0.636
Reliability (To assess how users of a DL perceive the trustworthiness of service provided)	6.39	0.641
Service quality (To assess in what ways and to what extent DL services satisfy users' needs)	6.36	0.688
User satisfaction (To assess in what ways and to what extent users are satisfied with services provided by a DL)	6.32	0.742
Usefulness (To assess in what ways and to what extent DL services are useful for users to achieve their tasks)	6.29	0.984
Responsiveness (To assess the reaction time to a user's request for a DL service)	6.21	0.834
Timeliness (To assess in what ways and to what extent services are offered to users in a timely manner)	6.11	0.736
Availability of DL staff (To assess in what ways and to what extent a user can easily contact staff of a DL for questions, feedback and comments)	5.89	0.997
Types of services (To assess the types of services provided by a DL)	5.82	0.801
FAQ/ Q&A (To assess whether and how many FAQs or Q&A a DL provides to help users)	5.46	1.122
Confidence (To assess in what ways and to what extent users have a positive attitude toward services offered by a DL)	5.43	1.299
Follow-up services (To assess in what ways and to what extent adequate and timely continuing services are provided to users by a DL when necessary)	5.43	1.118
User education (To assess the types of user education offered by a DL)	5.36	1.079
Types of unique services (To assess the unique types of services provided by a DL compared to other related DLs)	5.14	1.137
Customized services (To assess whether a DL offers personalized services based on user profile information or user requests)	4.89	1.055

〈Table 8〉 Importance of evaluation criteria in the dimension of user engagement

Criteria	Mean	STD
User feedback (To assess the types of user comments and suggestions received by a DL, and in what ways and to what extent these comments and suggestions are incorporated into the enhancement of the DL)	5.89	0.974
Resource use (To assess in what ways and to what extent users uses resources in a DL)	5.81	1.021
Help feature use (To assess which help features are offered to users, how frequently, and in what context users try to use help-related features in a DL)	5.79	0.834
Site visit (To assess how frequently users visit a DL website and the duration for each visit)	5.50	0.893
Integration with external applications (To assess in what ways users can export digital objects and integrate with external applications, such as slide presentation software)	5.50	1.209
User participation channels (To assess the types of channels available to users to communicate with the staff of a DL)	5.39	1.156
User knowledge contribution (To assess the ways users can contribute to DL content and organization through tagging, commenting, and adding their own objects)	5.00	0.935
E-commerce support (To assess the capabilities of ordering digital objects online)	4.89	1.506

〈Table 9〉 Importance of evaluation criteria in the dimension of context

Criteria	Mean	STD
Information ethics compliance (To assess whether a DL identifies and conforms to ethical issues related to DL creation and use)	6.61	0.636
Copyright (To assess whether a DL identifies and conforms to copyright issues)	6.25	0.792
Content sharing (To assess in what ways and to what extent stakeholders of a DL are willing to share their content)	6.00	1.192
Targeted user community (To assess in what ways and to what extent a DL engages in targeted user groups)	5.75	1.137
Collaboration (To assess in what ways and to what extent stakeholders of a DL work together; To assess in what ways and to what extent stakeholders of a DL cooperate with stakeholders in another DL)	5.75	0.953
Organizational mission (To assess in what ways and to what extent DL creation and use conform to organizational objectives)	5.43	1.221
Social impact (To assess in what ways and to what extent the use of a DL influences the society)	5.18	1.368

Finally, the subjects selected "information ethics compliance (6.61)", "copyright (6.25)", and "content sharing (6.00)" as most important evaluation criteria in the dimension of context. In particular, the subjects gave a comparatively higher score for "information ethics compliance" because they are academic users instead of general users. Following the top criteria, "targeted user community (5.75)" and "collaboration (5.75)" were tied by being ranked fourth. Comparatively speaking, "social impact (5.18)" and "organizational mission (5.43)" were considered less important in relation to context evaluation. This group of users cared more on rules and policies than the impact of digital libraries on society and organization. Table 9 presents the importance of evaluation criteria in the dimension of context.

3. Discussion and Conclusion

Identifying evaluation criteria is essential for the successful evaluation of digital libraries. Previous research suggested a variety of evaluation criteria in different dimensions of digital libraries. However, users' perspective has not been sufficiently investigated in the digital library evaluation framework. The present study investigated users' opinions on the importance of evaluation criteria for digital library evaluation. The unique contribution of this study lies in the comprehensive examination of users' perception of evaluation criteria across different dimensions of digital libraries. The ratings of evaluation criteria showed the most and least important criteria from users' perspective. Practically, the findings of this study assist library professionals in making their

decisions in regard to selecting appropriate evaluation criteria for different evaluation objectives.

Different stakeholders identify their DL evaluation criteria based on their needs, background, their own interests, and familiarity with DL concepts. Users rank higher the criteria related directly to use of collections, such as authority in collections, or metadata accuracy in information organization, or the quality of services. Their rankings reflect the expectations of digital library users. Users think less about the cost and efforts required for building DLs, which are of concern to digital librarians. The top criteria selected by users indicate what they care the most in digital library use. For example, in information organization, they rated accuracy, consistency and in-depth of metadata as the top criteria. Accuracy is, of course, essential for users to actually use the metadata for their research and learning, and in-depth of metadata is useful for them to obtain as much information as possible for each digital item. In developing digital libraries, digital librarians need to consider users' needs in regard to quality of collections, metadata and services. Interface design needs to offer multiple options for users to access the documents. At the same time, interface design needs to consider the special requirements from users with a variety of disabilities. Reliable and effective system performance is a key requirement for users to access digital libraries.

However, not all users are the same. The subjects of this study represent academic user group of digital libraries. In addition to general users' perceptions of digital library evaluation criteria, they also have their unique needs and opinions because of their academic background. In the dimension of effects on users, they ranked research productivity and learning effects as their top choices. Research and learning are the two academic goals for this group of users. The design of digital libraries in academic settings needs to put research and learning as the priority in terms of collection, metadata and interface design. For example, librarians need to work with instructors to determine the types of metadata needed for learning purposes in the development of related digital collections. In providing digital services, librarians have to come up with ideas that tailor to the needs and characteristics of academic users. In the dimension of context, subjects of this study chose "information ethics compliance", "copyright", and "content sharing" that are important to academics as the most important evaluation criteria. Digital libraries in academic settings need to provide information related to ethics compliance, copyright information and content sharing options to assure and guide users in using of digital items.

Certainly, this study has several limitations. The sample size might not be sufficient to represent a variety of users of digital libraries even though they are real users of digital libraries. The results from the analysis of thirty users, including faculty, undergraduate and graduate students, cannot be generalized to general public users. Also, the authors were not able to conduct a comparison analysis between different groups (e.g., faculty vs. students) due to the relatively small sample size for statistical tests. However, the findings of this study yield insightful information

on users' perceptions of digital library evaluation. In addition, the present study investigated only the user group among different stakeholders. Although the ultimate objective of digital libraries is to serve users, end-users do not have sufficient knowledge of DL administration, collection development, or preservation techniques. Therefore, the investigation of other expert groups, such as digital librarians and scholars, is imperative to develop a comprehensive evaluation framework. Further analysis is going to investigate to examine opinions from other stakeholders including scholars and digital librarians. In addition, the authors plan to compare the opinions of those three groups to identify similarities and differences among them.

References

Borgman, C. L., G. H. Leazer, A. J. Gilliland-Swetland, and R. Gazan. 2000. "Evaluating Digital Libraries for Teaching and Learning in Undergraduate Education: A Case Study of the Alexandria Digital Earth Prototype (ADEPT)." *Library Trends*, 49: 228-250.

Buttenfield, B. 1999. "Usability Evaluation of Digital Libraries." *Science & Technology Libraries*, 17(3/4): 39-50.

Candela, L., D. Castelli, P. Pagano,, C. Thanos, Y. Ioannidis, G. Koutrika, S. Ross, H. Schek, and H. Schuldt. 2007. Setting the Foundations of Digital Libraries: The DELOS Manifesto. *D-Lib Magazine*, 13(3/4). [cited 2014.1.14]. <http://www.dlib.org/dlib/march07/castelli/03castelli.html>.

Fox, E. A., R. M. Akscyn, R. K. Furuta, and J. J. Leggett. 1995. "Digital Libraries." *Communications of the ACM*, 38(4): 23-28.

Fuhr, N., P. Hansen, M. Mabe, A. Micsik, and I. Sølvberg. 2001. "Digital Libraries: A Generic Classification and Evaluation Scheme." *Lecture Notes in Computer Science*, 2163: 187-199.

Fuhr, N., G. Tsakonas, T. Aalberg, M. Agosti, P. Hansen, S. Kapidakis, C. Klas, L. Kovács, M. Landoni, A. Micsik, C. Papatheodorou, C. Peters, and I. Sølvberg. 2007. "Evaluation of Digital Libraries." *International Journal on Digital Libraries*, 8(1): 21-38.

Hill, L.L., R. Dolin, J. Frew, R.B. Kemp, M. Larsgaard, D.R. Montello, M.-A. Rae, and J. Simpson. 1997. "User Evaluation: Summary of the Methodologies and Results for the Alexandria Digital Library, University of California at Santa Barbara." *Proceedings of 60th ASIST Annual Meeting*. (pp. 225-243, 369). Medford, NJ: Information Today.

Hill, L. L., L. Carver, M. Larsgaard, R. Dolin, T. R. Smith, J. Frew, and M.-A. Rae. 2000. "Alexandria Digital Library: User Evaluation Studies and System Design." *Journal of the American Society for*

Information Science, 51: 246-259.

Jeng, J. 2005. "Usability Assessment of Academic Digital Libraries: Effectiveness, Efficiency, Satisfaction, and Learnability." *Libri*, 55: 96-121.

Joo, S. and J. Lee. 2011. "Measuring the Usability of Academic Digital Libraries: Instrument Development and Validation." *The Electronic Library*, 29(4): 523-537.

Marchionini, G., C. Plaisant, and A. Komlodi. 1998. " Interfaces and Tools for the Library of Congress National Digital Library Program." *Information Processing & Management*, 34(5): 535-555.

Matusiak, K. K. 2012. " Perceptions of Usability and Usefulness of Digital Libraries." *The International Journal of Humanities and Arts Computing*, 6(1-2): 133-147.

Nielson. J. 1993. *Usability Engineering*. Academic Press: Cambridge.

Noh, Y. 2010. "A Study on Developing Evaluation Criteria for Electronic Resources in Evaluation Indicators of Libraries." *Journal of Academic Librarianship*, 36(1): 41-52.

Saracevic, T. and L. Covi. 2000. "Challenges for Digital Library Evaluation." *Proceedings of the American Society for Information Science*, 37: 341-350.

Saracevic, T. 2004. Evaluation of Digital Libraries: An Overview. Presented at the DELOS Workshop on the Evaluation of Digital Libraries. [cited 2014.2.14]. -> need cited date.
<http://www.scils.rutgers.edu/~tefko/DL_evaluation_Delos.pdf>.

Tsakonas, G., S. Kapidakis, and C. Papatheodorou. 2004. Evaluation of User Interaction in Digital Libraries. In M. Agosti, N. Fuhr (eds.) Notes of the DELOS WP7 Workshop on the Evaluation of Digital Libraries, Padua, Italy.

Van House, N. A., M. H. Butler, V. Ogle, and L. Schiff. 1996. "User-centered Iterative Design for Digital Libraries: The Cypress Experience." *D-Lib Magazine*, 2.

Xie, I. 2006. "Evaluation of Digital Libraries: Criteria and Problems from Users' Perspectives." *Library & Information Science Research*, 28(3): 433-452.

Xie, I. 2008. "Users' Evaluation of Digital Libraries: Their Uses, Their Criteria, and Their Assessment." *Information Processing & Management*, 44(3): 1346-1373.

Ward, J.L. and S. Hiller. 2005. "Usability Testing, Interface Design, and Portals." *Journal of Library Administration*, 43(1): 155-171.

Zhang, Y. 2010. "Developing a Holistic Model for Digital Library Evaluation." *Journal of the American Society for Information Science and Technology*, 61(1): 88-110.