Chapter XXIX
Digital Library Requirements: A Questionnaire-Based Study

Stephen Kimani
Institute of Computer Science and Information Technology, JKUAT, Kenya

Emanuele Panizzi
University of Rome “La Sapienza,” Italy

Tiziana Catarci
University of Rome “La Sapienza,” Italy

Margerita Antona
FORTH-ICS, Greece

Abstract

The gathering of user requirements is key to the gaining of a deeper understanding of the needs evolving from the user’s operational context and from the use of the system. User requirements are pivotal in guiding the development process of any system. This is no less true in the arena of digital libraries (DLs). The gathering of DL requirements should be conducted with the understanding that the anticipated DL user interface should accord support to the user throughout the entire DL usage/interaction process. This chapter describes a questionnaire-based study of DL requirements based on the foregoing understanding. The study covered a wide range of issues pertinent to the design of user interfaces for DLs, including: user characteristics/profiles, current experience in DL usage, functional requirements, nonfunctional requirements, and contextual requirements. To the authors’ knowledge, this is the first systematic empirical investigation of DL requirements that covers such a wide range.
Digital Library Requirements

Introduction

It has been said that “the digital library in a broader context is nothing but a database ...the objective of any digitization process should be the empowerment of the people” (TERI, 2004). It is worth noting that there exist many definitions of the term digital library (e.g., Borgman, 1999; Digital Library Federation, 1999); however some of the common points across different definitions are: digital libraries (DLs) can be comprised of digital as well as nondigital entities; the realm of libraries is constituted not only of library objects but also of associated processes, actors, and communities; and the content of DLs can be extremely heterogeneous. It appears that the bottom-line DL issue in this matter is to provide a coherent view of the (possibly) large collection of the available materials (Lynch & Garcia-Molina, 1995). With the foregoing statement in mind we may, to a large extent, regard a DL as an information environment. In this environment, there are producers and consumers of information and knowledge (including collaborators and stakeholders), and they may change their role or play more than one role at the same time.

The gathering of user requirements is instrumental to the gaining of a deeper understanding of the needs evolving from the user’s operational environment and from the use of the system. User requirements are of key importance for guiding the development (and evaluation) process of any system. This is no less true in the arena of digital libraries. WP4 user interface and visualization is a cluster within the DELOS network of excellence on digital libraries. The cluster’s ultimate goal is to develop methodologies, techniques, and tools to establish a theoretically motivated and empirically supported frame of reference for designers and researchers in the field of user interfaces and visualization techniques for digital libraries, so as to enable future DL designers and developers to meet not only the technological, but also the user-oriented requirements in a balanced way.

During the first year, one of its specific objectives was to investigate the requirements for a DL user interface design. To that end, the cluster embarked on the gathering/identification of the functional requirements and nonfunctional requirements of DL users and stakeholders based on a questionnaire-based study.

The investigation involved a wide variety of issues pertinent to the design of user interfaces for DLs, ranging from user characteristics and profiles to current experience in the use of DLs, through functional and nonfunctional requirements, to requirements specifically related to the context of use. While there exist lots of related efforts (e.g., Marcum, 2003; Pasquinelli, 2002; Schilit, Price, & Golvchinsky, 1998), our study is, to the authors’ knowledge, the first systematic empirical investigation of DLs user requirements that covers such a wide range of aspects. In this chapter, we describe the overall requirements emanating from the study.

In the rest of the chapter, we describe the questionnaires that were used in the DL requirements study; we then give some background information pertaining to the questionnaire participants. The DL requirements are then described and we then highlight future trends and conclude the chapter.

User Questionnaires

Potential or real DL users (i.e., stakeholders and end users) have a key role toward the successful design of DLs, and the questionnaires can encourage them to word their recommendations on possible support offered by a DL. The questionnaires also enable DL users to contribute to the definition of the scope of DLs’ functional and nonfunctional requirements. Apart from their use in the context of the data collection performed by the cluster, the questionnaires can also be reused in the process of designing digital libraries for
Digital Library Requirements

the purposes of the user requirements collection phase. The questionnaires consist of four parts: user background and demographics; user’s current experience; DL functional requirements; and DL nonfunctional requirements.

Demographics and User Background

There were 45 library users (14 female, 25 male, 6 not specified) who responded to the online questionnaire. Most of the respondents range from 20 to 55 years old and they all come from Europe. Although many of them are in the field of computer science, the samples contain very divergent backgrounds (from computer scientists to humanities studies and librarians). Only a few participants reported a considerable degree of disability in one or more of the following categories: cognitive, intellectual, and visual. The sample of these 45 DL users is also characterized by multilingualism, high level of education, high experience with computing and the Internet, as well as relatively high experience with DLs. Questionnaire results indicate that the users frequently access a DL and thus they are aware of the weakness, advantages, and drawbacks of the current digital systems. In addition, as far as the type of access used for data retrieval is concerned, the vast majority of concerns are public or free access, indicating that the users are not willing to pay a lot for retrieving data and knowledge from a digital archive. Moreover, Web access is by far the most popular medium. Two thirds of the DLs gathered through the questionnaires support English as the only language of interaction. Slightly less than a quarter of all the DLs support both English and some other local language. Very few DLs offer multilingual support or only the local language.

Digital Library Requirements

This section describes the key needs and requirements concerning both the functionality of digital libraries and other nonfunctional characteristics related to interaction and importance of user interface design. The requirements are analyzed and then organized into high and low importance requirements for both DL stakeholders and end users. The realization constitutes a handy resource for guiding the development and evaluation of future DLs.

Functional Requirements

It is important for the developers and providers of DL services and systems to be well informed about the prevalent DL stakeholder and user functionalities. It is also important to know how important those functionalities are to the stakeholders and users. This subsection gives an analysis of various DL functional requirements in terms of prevalence and importance. The study analyzed the following functionalities: functions for locating information, functions for presenting resources, functions for personalization of content and services, facilities for communicating and collaborating with other DL users, and other common DL functions (e.g., social navigation support, multilingual support, personal annotation, notification/alerting services, glossaries, thesauri and dictionaries, printing/print preview facilities, and downloading/uploading facilities). In addition, some participants expressed their interest in, and attached importance to, the need for help and guidance.

• Integration of knowledge: Integration of knowledge is a stakeholder-oriented task. In fact, in the stakeholder’s questionnaire we listed the content management func-
tionalities of the DL. Among all the content management related functionalities listed within the questionnaire, deleting resources was rated as of low importance to stakeholders, whereas moderate importance ratings were allocated to: editing existing resources; creating new classification schemes; index facilities; creating new resources; retrieving content; and glossaries, thesauri, and dictionaries. The DL functions at the provider site that appear to be of high importance to stakeholders are: organizing resources; archiving resources; and storing metadata about resources (e.g., creator, content, technical requirements, etc.). Stakeholders rated locating resources, creating cross-reference links among similar resources, and storing metadata about resources as of highest importance. Figure 1 provides an overview of common DL functions (content management) at the stakeholder site and their importance to the stakeholders. The DL stakeholder functional requirements center around two main functional areas: content management and membership management.

- **Access to knowledge:** Among the functions for locating information, higher importance ratings were allocated to search (e.g., keywords search, parametric search), moderate importance ratings were allocated to index facilities and to navigation related functions (e.g., browsing predefined catalogues), and lower importance ratings were allocated to “See also” items (e.g., similar to the one at hand) and to functions for filtering search/browsing results (e.g., according to personal profile[s]).

- **Administration of content:** This is typically a stakeholder’s task. Among all the content management-related functionalities listed in the questionnaire, history facility, multilingual support, and bookmarks facility were of low importance to stakeholders. Moderate importance ratings were allocated to modifying existing classification schemes, retrieving services usage statistics, and updating end users on new/refined contents. Checking for inconsistencies appears to be of high importance to stakeholders. Stakeholders also expressed their interest/need for “access to control policies” and for retrieval of DL usage statistical data. The reader may refer to Figure 1 above.

- **Tool creation and management:** Among the functions for the personalization of content and services, higher importance ratings were allocated to functions for the presentation of contents according to profile, and to bookmarks facility (i.e., Favorites), moderate importance ratings were allocated
Digital Library Requirements

to the provision of suggestions for contents based on profile, and to services offered for profile definition (e.g., professional interests, personal interests, etc.), and lower importance ratings were allocated to the provision of suggestions for discussion with other library members with similar interest profiles, and to history facility. An overview of the ratings given by DL users is presented in Figure 2.

- **Membership management:** Among all the membership management-related functionalities listed in the stakeholder’s questionnaire, multilingual support, monitoring usage and identifying common patterns of use, and initiating special interest discussions were rated as of low importance to stakeholders. Moderate importance was allocated to updating end users on new/refined services and in maintaining a virtual help-desk for end users (e.g., FAQs facility). DL functions related to membership management that appear to be of high importance to stakeholders are services for subscribing and unsubscribing DL users.

- **Services – Interfaces for access to integrated knowledge:** Higher importance ratings were allocated to short description/previews, author(s)/editor(s), and title, moderate importance ratings were allocated to popularity (e.g., number of visits), to insertion/modification date, and to related items, and lower importance ratings were allocated to users’ ratings and to users’ discussions and reviews. An overview of the ratings is presented in Figure 3. As can be seen, the most important fields are “Title” and “Author.”

- **Services – Interfaces for sharing/integration of knowledge:** Among the functions for communicating and collaborating with other DL users, higher importance ratings were allocated to shared annotation facilities (e.g., peer reviews) and to e-mail services, moderate importance ratings were allocated to shared annotation facilities (e.g., peer reviews) and to e-mail services, moderate importance ratings were allocated to shared annotation facilities (e.g., peer reviews) and to e-mail services, and lower importance ratings were allocated to related items.

Figure 2. Common functions for personalization of DL content and services, and their importance to DL users

Figure 3. Common fields for presenting DL resources, and their importance to DL users
Digital Library Requirements to message boards services, and lower importance ratings were allocated to video conferencing and chat. Regarding communication, e-mail is the most frequent way required for communicating, while video conferencing appears not to be strongly desired. Furthermore, among all facilities for collaboration, “track changes” facilities received the lowest scores.

- **Other DL functionalities**: Higher importance ratings were allocated to printing/print preview facilities, and to downloading/uploading facilities, while moderate importance ratings were allocated to personal annotation, and notification (alerting) services, and lower importance ratings were allocated to social navigation support (e.g., through users’ rating of content) and to multilingual support.

**Nonfunctional Requirements**

This subsection gives an analysis of the importance of the following nonfunctional requirements: common usability requirements, common accessibility requirements, and other common nonfunctional requirements (e.g., ethics, safety, privacy, security, and system performance).

- General nonfunctional requirements at the user site and their importance according to DL users

In this part, the analysis of the importance of nonfunctional requirements according to the DL users is discussed.

**Usability Requirements**

Higher importance ratings were allocated to support for error prevention and handling, and ease of use of the DL, while moderate importance ratings were allocated to memorability and learnability, and lower importance ratings were allocated to subjective satisfaction users (i.e., expert, moderate, and novel users). An overview of the ratings is presented in Figure 4. The ease of use of a DL is the most important feature. On the other hand, the satisfactory performance for the experts is the least important one. This means that the users give high importance to the simple way of communicating, accessing, retrieving, and searching content of their interest. Furthermore, in terms of usability, some participants highly rated the importance of speed and flexibility to reach the information, and mentioned that the DL should “remember” the usability and accessibility preferences and requirements of the users.

**Accessibility Requirements**

Higher importance ratings were allocated to support for users with dexterity and mobility impairments, moderate importance ratings were allocated to support for users with speech, hearing, and learning impairments, and lower importance ratings were allocated to support for users with reading and visual impairments.

**Figure 4. Usability-related requirements and their importance to DL users**
Other Nonfunctional Requirements

Higher importance ratings were allocated to system performance, moderate importance ratings were allocated to security, privacy, and safety, and lower importance ratings were allocated to ethical requirements.

- General nonfunctional requirements at the provider site and their importance according to stakeholders

In this part of the chapter, the analysis of the importance of nonfunctional requirements is detailed according to the DL stakeholders.

Usability Requirements

Stakeholders expressed low interest in the memorability of the DL satisfaction to novel users, moderate interest in error prevention and handling and satisfaction to expert users, and high interest in the ease of use and of learning of the DL. Figure 5 provides an overview of the stakeholders’ preferences over specific usability attributes for DLs.

Accessibility Requirements

Stakeholders showed relatively high interest in special services and support for users with visual impairments, reading impairments, and mobility impairments, in comparison to special features for users with learning impairments, dexterity, speech impairments, and hearing impairments.

Other nonfunctional requirements

The stakeholders rated the other nonfunctional requirements as follows: high importance to safety, system’s performance (e.g., response time), privacy, scalability, and reliability (e.g., back-up facilities); moderate importance to interoperability (i.e., platform independence), extendibility (e.g., evolution of functionality and content-related requirements), adaptability (e.g., to different levels of user expertise), and compliance with national/international standards (e.g., ISO); and low importance to cost, interoperability (i.e., platform independence), portability, and ethics.

High and Low Importance Requirements

The foregoing specific functional and nonfunctional requirements were analyzed in order to identify high and low importance digital library requirements. In the sequel, the identified requirements are described, starting with the high and low importance requirements for stakeholders and then for end users.

stakeholder r equirements

Stakeholders appear to pay particular attention to functions for locating and organizing resources, including functions for creating cross-reference links among similar resources, and functions for
storing metadata about resources and checking for inconsistencies among the DL resources. Another interesting observation is that all accessibility requirements (i.e., for all kinds of disabilities) obtain a significant position in the list of requirements of high importance to DL stakeholders, whereas usability requirements fall in the group of moderately important features except the need for “ease of use” of the DL. In the group of requirements of high importance, DL stakeholders have also placed all kinds of functionalities related to the administration and management of registered DL users. On the other hand, the group of requirements that appear to be of lower value to DL stakeholders includes most of the miscellaneous functional and nonfunctional requirements, as well as usability requirements related to novel users. Furthermore, history facilities and multilingual support are of relatively low value to DL stakeholders.

**End-User Requirements**

DL end users, just like stakeholders, pay a lot of attention to all types of DL facilities for locating (subjectively) useful information. End users, in contrast to stakeholders, appear to pay particular attention to certain miscellaneous nonfunctional requirements such as system performance, security, privacy, safety, and other ethical requirements. Printing and up/downloading facilities are also of significant importance to DL users, followed by general usability requirements and accessibility for people with motor impairments (i.e., mobility and dexterity impairments). On the other hand, personalization is not really an important issue among DL end users, and facilities for user-to-user communication and collaboration are almost not an issue at all.

**Main Lessons**

The results of this initial work have led to the recognition and realization of some significant dimensions of diversity in users issues for DLs. Concerning demographics and user characteristics, the performed investigation has established that users from a number of European countries exhibit a variety of characteristics concerning age, educational level and professional background, and purpose of DL use. Nevertheless, most users have a recurrent DL usage pattern, that is, twice a week for a duration of 1 to 5 hours. Additionally, users clearly show a tendency to use free access DLs through the World Wide Web. Functional requirements and nonfunctional requirements vary according to the users’ role in DLs usage. In particular, it has emerged that end users and stakeholders of DLs tend to have different views and conceptualizations of DL systems and of their use. Overall, it appears that DL stakeholders strive for enriched functionality, whereas DL users pay more attention to the perceived behavior and reliability of a DL. At the present stage, a strong need for communication or collaboration in the DL end-user community does not seem to have emerged yet. End users view DLs as personalized environments where privacy is protected. On the other hand, stakeholders appear to view DLs as more collaborative environments. From the acquired data, it also emerges that the traditional “paper document” metaphor is still felt as prevailing in DL environments, based on the high degrees of importance placed on retrieval, printing, uploading, and so forth of documents, and with respect to the low importance placed, for example, on more digitally-oriented functions such as personal annotation of digital data. A requirement common to both end users and stakeholders concerns help and guidance. On the providers’ side, appropriate information about all the functionalities and the facilities of a DL are necessary in order to be able to provide the users with satisfactory results. While it is somewhat natural that end users and stakeholders focus on different functions of a DL, it appears that they also view nonfunctional requirements differently due to their different roles. In particular, it appears that they have different views on accessibility and
usability issues. The end users appear to have specific expectations concerning the reliability
and usability of the DL, whereas stakeholders
appear to be more aware of accessibility issues,
and to pay attention to the provision of DL ac-
cess to diverse target user groups. Stakeholders,
as professional and expert users, also consider
as more important the effectiveness of a DL user
interface with respect to its ease of use, whereas
end users, who may use a DL on a more occasional
basis, appear to be more interested in issues such
as usability, reliability, and safety. Diversity also
appears when analyzing users’ information needs
when searching DLs. In searching behavior, users
develop different individual styles, or employ
combinations of methods to establish relevance
judgments. Additionally, users, as opposed to
stakeholders, do not appear to have a clear under-
standing of other possible purposes of DLs use
than searching for information or document.

The above differences may find an explanation
in the different levels of awareness, among end
users and stakeholders, concerning the techno-
logical aspects of DLs that can differentiate them
from paper-based libraries, by providing access,
for example, to non-textual material, as well as
support for information exchange, annotation,
and cooperation.

The above shed light on some important aspects
of the users’ access to DLs. Such aspects need to
be carefully considered in order to further inves-
tigate the implications on the design of DLs user
interfaces. For example, given the differences in
the priorities between end users and stakeholders,
it is unlikely that a single user interface can satisfy
the requirements of both target groups. This needs
to be reflected also in appropriate (possibly differ-
ent) metaphors, carefully elaborated in order to go
beyond the physical library and paper-based model
towards a more novel concept of a digital library
as a collection of diverse (possibly multimodal)
digital resources (e.g., software, services, etc.) in
desktop and mobile environments.

Future Trends and Conclusions

We are living in a day when people are increasingly
using a variety of computing devices in their daily
lives which are not all continuously connected
to a network. Such devices present an interesting
opportunity for the creation of personalized
information spaces, such as, DLs with collections
and services that correspond to targeted needs and
situations. It is also worth stressing that DL users
are changing their way of using DLs. DL users are
nowadays often playing more and more different
roles at different times and places, for example,
as consumers of information and producers of
information at the same time (Bertini, Catarci,
Di Bello, & Kimani, 2005).

The prioritization of requirements elaborated
in this chapter has the potential to lead to the
elaboration of a framework for DL user interface
design. A future research consideration would be
to increase the sample size of both the end user and
stakeholder questionnaires. This would facilitate
a richer understanding of the causal nature that
exists between the respondent, functional require-
ments, as well as the nonfunctional requirements.
It would also make it possible to use the results
to formulate population inferences. Another
important issue to be addressed in the context
outlined above concerns the DL lifecycle. The
investigation of the DL lifecycle will be targeted
to provide an insight into how a DL is expected to
evolve regarding users and stakeholders interac-
tion, and how the different phases of the lifecycle
are interrelated with functional or nonfunctional
requirements.

This chapter has described the overall require-
ments emanating from the questionnaire-based
study. It has in particular described the functional
and nonfunctional requirements that were ob-
tained from the study. Moreover, the chapter has
identified and reported the high and low impor-
tance DL requirements based on an analysis of the
Digital Library Requirements

foregoing specific functional and nonfunctional requirements. The chapter has also discussed the main lessons that were learned from the study.

key Terms

**Demographics:** Demographics refer to the characteristics of human populations and population segments. Demographics include aspects such as age, gender, nationality, occupation, language, and so forth.

**Digital Library:** Borgman (1999) defines digital libraries as:

A set of electronic resources and associated technical capabilities for creating, searching and using information. In this sense they are an extension and enhancement of information storage and retrieval systems that manipulate digital data in any medium (text, images, sounds; static or dynamic images) and exist in distributed networks. The content of digital libraries includes data, metadata that describe various aspects of the data (e.g., representation, creator, owner, reproduction rights), and metadata that consist of links or relationships to other data or metadata, whether internal or external to the digital library.

According to the Digital Library Federation (Digital Library Federation, 1999):

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.

**Functional Requirements:** They describe system services or functions.

**Nonfunctional Requirements:** They define a constraint on the system or on the development process.

**Questionnaire:** A questionnaire is a written list of questions that is distributed to subjects/participants to gather information.

**Stakeholders:** Stakeholders are people who do not use the system directly but they may be affected by the system.
**Digital Library Requirements**

**Users:** Users are people who use/interact with the system directly.

1. http://www.delos.info
2. The reader may refer to the cluster Web site http://delos.dis.uniroma1.it for the full and more detailed report from the study.

3. The questionnaires have been made available to the public in a Web-based form, and can be accessed from the cluster Web site http://delos.dis.uniroma1.it.

4. Please note that in all figures with charts, the items in the legend are ordered such that they have a direct correspondence with the bar items in the graph itself, for example, the top-most item in the legend corresponds to the top-most bar item in the graph.