



Exploring usefulness and usability in the evaluation of open access digital libraries

Giannis Tsakonas ^{*}, Christos Papatheodorou

Laboratory on Digital Libraries and Electronic Publishing, Department of Archive and Library Sciences, Ionian University, Palaia Anaktora, Plateia Eleftherias, 49100 Corfu, Greece

Received 31 January 2007; received in revised form 12 July 2007; accepted 13 July 2007
Available online 4 September 2007

Abstract

Advances in the publishing world have emerged new models of digital library development. Open access publishing modes are expanding their presence and realize the digital library idea in various means. While user-centered evaluation of digital libraries has drawn considerable attention during the last years, these systems are currently viewed from the publishing, economic and scientometric perspectives. The present study explores the concepts of usefulness and usability in the evaluation of an e-print archive. The results demonstrate that several attributes of usefulness, such as the level and the relevance of information, and usability, such as easiness of use and learnability, as well as functionalities commonly met in these systems, affect user interaction and satisfaction.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Digital libraries evaluation; Open access; e-Prints; Interaction; Usefulness; Usability

1. Introduction

The various types of engagement in digital libraries (DLs), such as funding developmental stages or devoting important shares in the acquisition of access has drawn considerable attention to the evaluation of these systems during the late period. The term digital library is vast, covers many and different applications and has been used interchangeably for systems, like digitized collections, e-journals platforms, network databases, library websites, etc. Moreover the current electronic publishing business models enrich the DL technology aiming to provide powerful information access options to the users. However despite the polymorphous nature of DLs, one feature is stable: information provision through networked systems.

The advance of open access systems increases the complexity and diversity of DLs because the collection development requires the involvement of new actors; e.g. authors submit their own original material, system editors check and approve the corresponding metadata, etc. Although a significant amount of research is noted on the evaluation of commercial or research DLs, the equivalent work is not dedicated to the freely

^{*} Corresponding author. Tel.: +30 2610 969625; fax: +30 2610 969672.

E-mail addresses: john@lis.upatras.gr (G. Tsakonas), papatheodor@ionio.gr (C. Papatheodorou).

accessible DLs. However the evaluation of such systems is imperative, due to their worldwide acceptability and usage, but their special characteristics perplex the whole process of evaluation. Different types of evaluation are proposed and various approaches are followed. Their majority is concentrated on the intrusion of these systems to the scholar population and the effect of the social, economic and technologic environment on their embracement.

In the present paper the prime investigation aim is to define which content and system features most significantly affect the usefulness and usability levels of an Open Access (OA) repository. In particular a user-centered evaluation approach is presented and applied to the subject repository E-LIS (<http://eprints.rclis.org>), which is an e-print archive in the field of library and information science. This investigation is done by applying a DL evaluation framework – with a special interest on usability and usefulness – on important attributes and functionalities of E-LIS.

Section 2 resumes the scientific literature in regard to evaluation of DLs and OA systems. Section 3 presents the research setting in terms of problem, aims and models employed. Section 4 presents the methodology of the current study, while Section 5 presents the results. Section 6 is dedicated to the discussion of the main findings and Section 7 summarizes and suggests further research steps.

2. Literature review

There are various DL evaluation initiatives that use different protocols and models; some of them continuously tested, other created or adapted to fit to specific systems and needs. Among the various concepts under investigation, two have gained significant attention in user-centered evaluation, namely usefulness and usability. Originating from the fields of information behaviour and human–computer interaction respectively, they both hold significant role in pursue of user satisfaction and system usage, as they study users' interaction with representations of information systems and information objects. While the information behaviour domain concentrates on the dynamic expression of user activities in a given context and may identify areas in system usage that require improvement, human–computer interaction aims – through user-centered approaches – at the development of systems that enhance user performance and satisfaction.

Järvelin and Ingwersen (2004) have based the formation and analysis of user's cognitive processes and information seeking behaviour on the interactivity between the user and information resources structural elements (e.g. systems, items and interfaces). The effect of information sources structure have been acknowledged in the information behaviour literature as one of the many intervening variables in the information seeking process, that can be "*supportive of information use, as well as preventive*" (Wilson, 1999).

The user interaction with content and system has been used in the area of information architecture for the purposes of effective DL design. Toms (2002) has used this tri-polar structure as a platform, where users' information seeking behaviour takes place, and upon this platform proposes ways of extracting meaningful design suggestions. Several empirical studies have demonstrated that usability problems, i.e. design inefficiencies, can disrupt user's information searching activity and affect information task accomplishment. Kim (2002) has employed DLISP model to match information searching expressions with usability problems, while Stelmaszewska and Blandford (2002) have demonstrated the need for effective usability design in DLs for assisting the creation of information searching strategies. Current studies in the DL area (Chowdhury, Landoni, & Gibb, 2006) show that there is a consensus in the research community for unified treatment of the two concepts.

Additionally, the interest on the concepts of DL usefulness and usability is reinforced by the opinions of the users themselves. Xie (2006) collected user identified DL evaluation criteria, analyzed and classified them in five categories, namely usability, collection quality, service quality, system performance efficiency and users opinions. The principal finding was that "*users like to apply the least effort principle to finding useful information to solve their problems*". Results of a similar study by Kani-Zabihi, Ghinea, and Chen (2006) designate that users prefer learnable and reliable DLs, than aesthetically pleasant and supportive ones. The authors of the study examined the users' preferences on system functionalities, interface usability and content and found that learnability and reliability were prerequisites for DL satisfaction for all types of participants (from novices to experts).

Similarly the two concepts have been extendedly discussed in the field of information systems (IS) management. Technology Acceptance Model (TAM) is widely used in the field of IS success and it has been used in

prior studies of predictive usage of digital libraries (Hong, Thong, Hong, & Tam, 2002; Thong, Hong, & Tam, 2002), search engines (Liaw & Huang, 2003) and portals (Yang, Cai, Zhou, & Zhou, 2004). TAM exploits users' perceptions of usefulness and ease of use of an information system, relates them with external variables of users' profile and combines them to predict attitudes and actual usage. According to the Delone and McLean (1992), one of the eminent models in IS success, system and information quality affect seriously user satisfaction. Therefore the qualitative properties of the system and the information provided can improve system's use and consequently affect the organizational performance.

Concerning the DL evaluation methodology Bertot, Snead, Jaeger, and McClure (2006) report an iterative campaign based on the system-driven concepts of functionality, usability and accessibility. One of the main suggestions by the authors is that evaluation activities – in order to assist the formation of a holistic picture for user interactivity and to extract easier a set of implications on system design – should be based on a number of constructs (evaluation concepts), rather than to focus on one. Accordingly, DigiQUAL (Kyrillidou & Giersch, 2005), which has its origins in the field of service quality in the library sector and aims to be extended to the digital space of service provision, covers areas such as design features, accessibility/navigability, interoperability, DL as community for users, developers and reviewers, collection building, role of federations, copyright, resource use, evaluating collections and DL sustainability. It is quite evident that the level of focus depends on several parameters, such as the purpose of the research, the macroscopic view of the evaluation, the object of the evaluation, etc. Saracevic (2004), in a meta-analysis study of the current evaluation practices, has classified the various parameters in four main categories, namely construct, context, criteria and methodology. This classification supports any description endeavour and differentiates each evaluation initiative for the sake of meaningful comparison. The need for identification of such parameters is recognized by the DL evaluation community (Fuhr, Hansen, Mabe, Micsik, & Sølvberg, 2002; Fuhr et al., in press), by finding integral components of the DL operation (such as content, system and user) and matching them to criteria, metrics and methods.

Since the present paper focuses on OA systems some representative studies in this particular field are presented. Although open access DLs and electronic information services have undergone various evaluations of their impact, their technological infrastructure (Wyles, Maxwell, & Yamog, 2006) and integration with library resources (Center for Research Libraries, 2006), there are a few studies concerning the interaction of the users with such systems. By the number of user studies it can be figured out that the acceptability of these systems and the corresponding services by the scientific community is an attractive research issue. Nicholas, Huntington, and Rowlands (2005) gathered the opinions of a world wide sample of scientific authors to estimate the rate of adoption of OA publishing schemas and to compare it with the dominant publishing habitus. They showed that scholars from disciplines that have dominant subject repositories, like physicists, computer scientists and economists, relate OA with the notion of “good access” to electronic resources. Kurata, Matsubayashi, Mine, Muranushi, and Ueda (2006) showed that the usage of e-prints servers in Japan is limited to specific disciplines, such as physicists, and that electronic journals remain the principle vehicle of scientific communication. Their results also show that the highest levels of preference for alternative publishing means, such as e-prints, are traced in disciplines that have strong backgrounds in subject-based repositories, such as physics (which is supported by the Arxiv repository, <http://www.arxiv.org>). This is supported by several studies supporting the argument that scholars in various disciplines are favourable to self-archiving, due to the existence of dominant subject repositories (Swan & Brown, 2005). However Andrew (2003) and Davis and Connolly (2007) mention that, unlike scholars from other disciplines, physicists in their respective studies do not self-archive their research outcome in institutional repositories, because of the existence of Arxiv. Despite all these conflicts, one definite outcome of the various studies is that scholars from disciplines that facilitate subject repositories are aware of the self-archiving processes, while as an explanation of this paradox can be proposed the type of OA vehicle (e.g. repository, journal) in which they self-archive and the associated rewards.

Digital repositories, either subject-based, or institutional, conjoin with two different types of usage: archiving, or input activity according to Luce (2001), and actual usage. While there are studies examining the use of digital repositories (Correia & Castro Neto, 2002), the major part of these studies show that the self-archiving aspect of these systems' use is associated with publication aspects (e.g. visibility, preservation and durability of work), communication and self-related benefits (e.g. funding and curriculum strengthening) (Kim, 2005; Kim,

2006; Rowlands & Nicholas, 2005). The literature also shows that the interest in OA systems is concentrated on electronic journals and institutional repositories, presumably because of the association of these repository variations with the previously mentioned benefits. This makes easier the application of various means for the growth of collection and the improvement in quality terms of the repositories, such as mandating policies or peer-review processes, but it is not easy to apply in the cases of subject-oriented repositories. The above studies also suggest that the acceptability and the adoption of OA e-print systems (in their various formulations) is related to contextual factors, strongly linked with traditional structures and mechanisms of the scholar environment, and require the development of several synergies between scholars, libraries, funding agencies, government and non-government organizations and other agents.

Hitchcock et al. (2002) evaluated the usefulness and usability of the open-access bibliographic service Citebase. Although this formative evaluation campaign demonstrated a useful and usable system, problems with the coverage and the navigation ability were reported. Despite the fact that Citebase is not a DL, but a bibliographic service that in parallel provides access to full text, where available, the results showed problems concerning the features of content and system. Silva, Laender, and Gonçalves (2007) examined the various aspects of input activity in the BDBComp self-archiving system. In their usability study they recorded the average submission time under different conditions, such as users' experience, expertise and content types, and gathered their opinions on easiness to use the e-print archive, comfort and usefulness. On the other hand, Kim (2005) investigated several aspects of the usage activity of two digital repositories systems, namely DSpace and Eprints, in the context of Australian National University institutional repository. He went on to count the time demands of the systems, their error proneness and user satisfaction. Concluding the evaluation activities of OA systems, beyond issues commonly used in the evaluation of DLs (such as multi-constructness), involves also the examination of aspects related to their purpose and their typical user tasks' supporting processes.

3. Research setting

As literature suggests, user-centered evaluation of a DL should be multi-constructed, capable to capture a panoramic view of users' opinions, able to take into account their characteristics and grounded on their perceptions and goals. Thus for the aims of this research Interaction Triptych Framework was used (ITF) (Fuhr et al., *in press*; Tsakonas, Kapidakis, & Papatheodorou, 2004), which is a theoretical model that attempts to integrate knowledge and experience from the fields of information behaviour and human computer interaction. ITF is based on the concept of DL components interaction. Each DL is consisted by three main components, namely system, content and user. Each component interacts with the other two and each interaction defines an evaluation axis, which is considered as the resultant of a set of descriptive attributes of that interaction. In particular the interactions between the three components (i.e. the evaluation axes) are defined as usability, usefulness and performance and each of them is a resultant of a number of attributes, which are considered as evaluation variables (Fig. 1). ITF complies with the requirements extracted by the previous experience as (a) it fulfils the multi-construct requirement since it has the ability to explore the two main categories of user-centered interest in DL evaluation, (b) it provides quick and convenient view of the users' opinions on critical factors of the evaluated system, allowing further exploration with different methodologies, and (c) it takes into consideration important variables of the users' profile and attitude. The following subsections present analytically the evaluation axes and their components attributes.

3.1. Usefulness

The concept of usefulness defines whether DLs constitute valuable tools for the completion of users' tasks. Usefulness answers the questions if DLs support users' information needs and work completion. Users' work tasks are formed by their social and organizational context and responds to needs like research, authorship, etc. Users seek in DLs the appropriate resources for their work tasks, both relevant, in terms of subject proximity, and "integrateable", in terms of content morphology. On the other hand information tasks include all actions related to need formulation, need expression, querying, relevance assessment, all combined and executed in an iterative manner. During all those phases of information seeking activity there are common attributes that are expressed either as requirements or criteria, depending on the stage of the information seeking

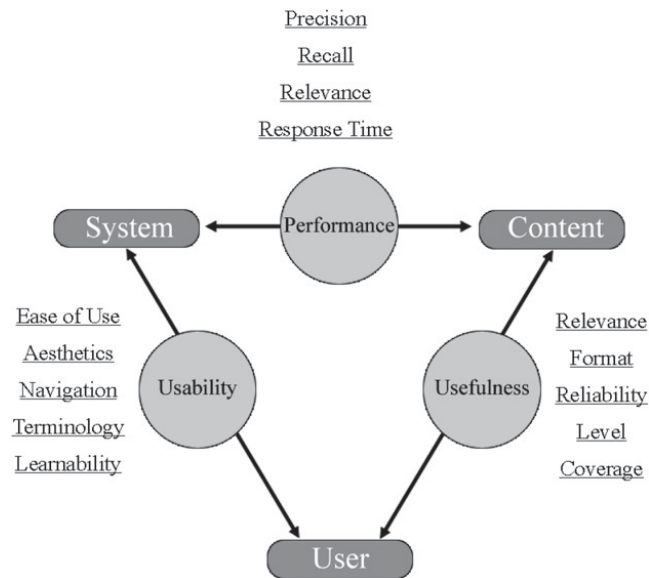


Fig. 1. Interaction triptych framework.

procedure. ITF summarizes these attributes into relevance, reliability, level of the provided information, format and temporal coverage.

Hong et al. (2002) report that system features have effect on both users' perceptions of ease of use and usefulness. However relevance, which is considered as a system feature, was related only to perceived usefulness. The authors interpret this relation as the association of relevance to the content of the DL. Liu (2004) signifies the importance of information reliability and credibility for the selection of the appropriate resources, while Vakkari and Hakala (2000) include in their relevance criteria the level of the provided information. Moreover users' information searching behaviour has demonstrated that despite retrieval of full text resources is significant, other levels of information, such as abstracts, are also preferred (Wolfram & Xie, 2002), probably due to the content overview they provide to users (Krottmaier, 2002). Current evaluation practices involve recording and analysis of usage statistics based on users' format preferences, such as .pdf and .html file formats (Mercer, 2000). Xie (2004) includes content coverage among the important components of an information retrieval system and associates it with other attributes of the information resources, such as reliability and level, while the library community (Brennan, Burkhardt, McMullen, & Wallace, 1999) jointly assesses the ratio of coverage and level in pursue of collection quality.

3.2. Usability

Usability stands on the user-system axis, focuses on the effective, efficient and satisfactory task accomplishment and aims to support a normal and uninterrupted interaction between the user and the system. DL community has shown an increasing interest in usability and through the research activities a set of attributes have been identified, such as ease of use, terminology, navigation, aesthetic appearance, learnability. Easiness of use is considered as a crucial attribute of DL interaction, especially in advanced systems, like aggregated search interfaces (Park, 2000). Previous usability studies (Ebenezer, 2003; McMullen, 2001) have shown that terminology raises important barriers in user's understanding of principal functions and contribute to negative changes in their affective state. Navigation is considered as an important factor contributing to the improvement of users' performance, as they are able to trace their place in the DL and to direct to previous or next destinations (Hartson, Shivakumar, & Pérez-Quinones, 2004; Theng, Duncker, Mohd-Nasir, Buchanan, & Thimbleby, 1999). Additionally the aesthetic appearance and layout has a crucial role to the overall satisfaction rate (Jackson, 2001). Van House (1996) suggested a simplified interface that would reduce users' efforts and recent usability studies have concentrated on the effect of inappropriate visual layout to user interaction

(Allen, 2002; Cockrell & Jayne, 2002; Fuller & Hinegardner, 2001). Finally some researchers have emphasized on the ability of a DL to be easy to learn in order to improve user familiarity and performance (Ferreira & Pithan, 2005; Jeng, 2005; Sutcliffe, Ennis, & Hu, 2000).

3.3. *Performance*

Performance evaluation is often a system-centered process, based on quantitative data and does not involve real users. Precision (the ratio of relevant items that were retrieved to a set of retrieved items) and recall (the ratio of relevant items retrieved to a set of relevant items of a collection) are the main metrics, which are transferred from the field of Information Retrieval. With the advance of networked systems other metrics are introduced, such as response time (Kobayashi & Takeda, 2000). Response time is considered as a metric sensitive to the subjective judgement of users and is often reported by end users as a criterion for the adoption of a DL. The proliferation of mechanisms for the retrieval from the internet and the empowerment of the users have initiated a discussion about the extension of IR metrics in order to cover new aspects, such as the search engine design and functionalities (Landoni & Bell, 2000).

4. Methodology

4.1. *Aim of the research*

The present study seeks to find which DL interaction attributes affect most significantly E-LIS usefulness and usability. This is obtained using two tools: (a) following an online questionnaire survey and inferential statistical analysis to define these attributes, and (b) using regression analysis to identify their predictive strength, as well as the impact of the system's functionalities.

4.2. *Evaluated system – E-LIS*

According to the repositories typology proposed by Heery and Anderson (2005), E-LIS is an international e-print archive that holds full text pre-print and post-print documents and their metadata and provides enhanced access to researchers in the domain of field of librarianship and information science. E-LIS' functionalities provide: free access to content through simple and advanced search interfaces and various indices, personalized access for document submission or review and approval (according to user roles and rights), and several other tools and peripheral services, such as document interlinking or email-based awareness. E-LIS is an effort that covers more than 80 countries and hosts more than 6000 documents in many different languages (as of July 9, 2007). The main requirements for the inclusion of archives in E-LIS are domain relevance and document integrity for the promotion of scientific communication. E-LIS is grounded on a solid and vivid international community, which has structured communication channels and organized dissemination events for the achievement of internal consistency. In each country-member there are one or more editors who are responsible for the promotion of E-LIS, the support of native users and the control of the deposited archives.

4.3. *Data collection procedure and instrument*

Data for this survey were gathered by an online questionnaire. The selection of this procedure was suggested due to the geographical dispersion of the E-LIS community, the time saving aspect and the ability of the instrument to gather in a quantitative manner the users' opinions and satisfaction (Covey, 2002; Allen, 2005). A call for participation was distributed to several national and international mailing lists by the respective national E-LIS editors. Editors were also invited to take part in the survey, as well as the administrative personnel. The call was also visible on the main E-LIS website and the Greek website for a period of one month (May to June 2006). A total of 131 of valid questionnaires were collected for analysis, after the removal of duplicate and test submissions.

The questionnaire (<http://dlib.ionio.gr/~gtsak/e-lis>) consisted of thirty four (34) questions, structured in three main parts.

Table 1
Cronbach α

Usefulness		Usability		Performance	
Group	.910	Group	.939	Group	.854
1.1	.902	2.1	.922	3.1	.795
1.2	.895	2.2	.932	3.2	.858
1.3	.886	2.3	.924	3.3	.823
1.4	.886	2.4	.929	3.4	.775
1.5	.906	2.5	.927		
1.6	.888	2.6	.931		

The first part (six questions) consisted of questions that asked users about their familiarity with E-LIS, the perceived usage, the attributed significance of E-LIS in their information seeking activity and the potential dedication of time and effort to retrieve the desired information. The second and more important part (16 questions) investigated the role of the usefulness, usability and performance attributes on the evaluation of E-LIS DL interaction. In detail, it consisted of three subsets of questions each one corresponding to the concept of usefulness, usability and performance respectively. Each subset included also a final question measuring the overall satisfaction with the related concept. The last part of the questionnaire (12 questions) examined the relative influence of the E-LIS functionalities on the concepts of usefulness and usability. In all questionnaire parts, participants were invited to deposit their degree of agreement to a number of statements through a five-point Likert scale (from “Disagree” to “Agree”).

Reliability of the research instrument was measured. Cronbach alpha tests were conducted separately for each group of items and rates were high for every group and every separate item, respectively (Table 1), superseding the 0.70 threshold proposed by Nunnally and Bernstein (1994). There was an exception in the question 3.2 (Response Time), which scored higher than the value of its group, but the difference was not considered that affected drastically the group's value.

5. Results

5.1. Characteristics of the participants

Table 2 presents the respondents' relation with the archive, such as role, geographic area (continent-wide) and E-LIS' introduction channel. The results show that almost half of the participants were unregistered users, indicating an equal readership degree of E-LIS compared to its active membership. The results also show that E-LIS has a plenitude of dissemination channels, which introduce the service to the potential users. It is noticeable that almost half of the respondents learned about E-LIS by informal means, like their own searches in the Internet or personal communication with friends and peers.

Table 3 presents the results of the questions reflecting users' characteristics. Participants reported their usage frequency, appointed significance and intention to spent time and effort to complete successfully their work tasks through E-LIS. Although the average usage of the DL is below the mean value of the users' rates, there is a normal distribution recorded that indicates several types of usage, from very frequent to infrequent.

Table 2
Demographic data

Role	Area		Introduction channel		
	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	
Editors	17 (12.98)	Africa	2 (1.53)	Announcements	29 (22.14)
Registered users	44 (33.59)	Asia	14 (10.69)	By my own	31 (23.66)
Unregistered users	66 (50.38)	Europe	76 (58.01)	Papers	20 (15.27)
Other	4 (3.05)	N. America/Caribbean	22 (16.79)	Personal communication	33 (25.19)
		S. America	12 (9.16)	Other	18 (13.74)
		Oceania	5 (3.82)		
Total	131 (100.00)		131 (100.00)		131 (100.00)

Table 3
Descriptive statistics (1 = disagree, 5 = agree)

	1 <i>n</i> (%)	2 <i>n</i> (%)	3 <i>n</i> (%)	4 <i>n</i> (%)	5 <i>n</i> (%)	\bar{x}	SD
Usage	24 (18.32)	24 (18.32)	38 (29.01)	20 (15.27)	25 (19.08)	2.98	1.359
E-LIS Significance	21(16.03)	32 (24.43)	45 (34.35)	20 (15.27)	13 (9.92)	2.79	1.183
Time	7 (5.34)	18 (13.74)	43 (32.82)	31 (23.66)	32 (24.43)	3.48	1.159
Effort	7 (5.34)	18 (13.74)	32 (24.43)	39 (29.77)	35 (26.72)	3.59	1.176

Accordingly, participants appoint mediocre importance to E-LIS in their regular information seeking activity, while they are more positive towards the statements of spending significant amounts of time and effort to complete their information tasks.

Cross-tabulation of the E-LIS role and the usage revealed that usage is consistently increased when the type of respondent indicates an active role, such as editor (Agree $n = 10$) or Registered User (Agree $n = 11$). Likewise E-LIS significance and intention to spent the needed time and effort for the retrieval of the desired information increase as the participants report more intense usage (Agree $n = 10$ in all cases). Analysis of Variance reported that the factor of E-LIS' role shows considerable differences at the questions of usage ($F = 9.016$, $p > .001$) and appointed significance ($F = 3.466$, $p > .010$). However there are non-significant differences at the questions for time ($F = 0.984$, $p > .419$) and effort spent ($F = 1.059$, $p > .380$). The same pattern was observed at the reverse cross-tabulation where the factor usage showed significant differences with role ($F = 13.913$, $p > .001$) and appointed significance ($F = 40.399$, $p > .001$), but did not show with time ($F = 1.099$, $p > .360$) and effort ($F = 0.437$, $p > .782$). Therefore the rest of this section presents and analyses the results in terms of role and usage factors. This analysis is followed because (a) usage is considered a critical factor for creating and classifying user profiles, as it implies experience in effective and efficient use of DLs and (b) the role in the archive is introduced by the system's nature and accordingly suggests some sort of experience with the related processes (input/usage activities).

5.2. Results on axes and features

This section reports the findings of the second part of the questionnaire, which focused on the significance of usefulness, usability and performance attributes. Table 4 presents the participants' answers, which demonstrates in general a positive attitude towards each attribute.

Table 4
Descriptive statistics for questions in features categories (1 = disagree, 5 = agree)

	1 <i>n</i> (%)	2 <i>n</i> (%)	3 <i>n</i> (%)	4 <i>n</i> (%)	5 <i>n</i> (%)	\bar{x}	SD
<i>Usefulness</i>							
1.1. Relevance	8 (6.11)	9 (6.87)	30 (22.90)	47 (35.88)	37 (28.24)	3.73	1.129
1.2. Format	5 (3.82)	8 (6.11)	32 (24.43)	42 (32.06)	44 (33.59)	3.85	1.075
1.3. Reliability	8 (6.11)	7 (5.34)	34 (25.95)	55 (41.98)	27 (20.61)	3.66	1.058
1.4. Level	6 (4.58)	14 (10.69)	25 (19.08)	43 (32.82)	43 (32.82)	3.79	1.150
1.5. Coverage	7 (5.34)	16 (12.21)	47 (35.88)	36 (27.48)	25 (19.08)	3.43	1.096
<i>Usability</i>							
2.1. Ease of use	5 (3.82)	3 (2.29)	25 (19.08)	44 (33.59)	54 (41.22)	4.06	1.021
2.2. Aesthetic	5 (3.82)	8 (6.11)	39 (29.77)	48 (36.64)	31 (23.66)	3.70	1.021
2.3. Navigation	6 (4.58)	6 (4.58)	23 (17.56)	54 (41.22)	42 (32.06)	3.92	1.045
2.4. Terminology	4 (3.05)	7 (5.34)	21 (16.03)	56 (42.75)	43 (32.82)	3.97	0.992
2.5. Learnability	4 (3.05)	3 (2.29)	23 (17.56)	50 (38.17)	51 (38.93)	4.08	0.966
<i>Performance</i>							
3.1. Precision	8 (6.11)	22 (16.79)	48 (36.64)	39 (29.77)	14 (10.69)	3.22	1.047
3.2. Response Time	4 (3.05)	5 (3.82)	31 (23.66)	48 (36.64)	43 (32.82)	3.92	0.997
3.3. Recall	6 (4.58)	20 (15.27)	54 (41.22)	40 (30.53)	11 (8.40)	3.23	0.965

5.2.1. Usefulness questions

All of the usefulness features (1.1–1.5) scored high rates, indicating a very positive attitude towards the usefulness attributes. However, the only feature that failed to clearly satisfy users is Coverage. It seems that the participants were not satisfied entirely with the content coverage of E-LIS. The comparison of responses between different roles demonstrated that participants with more active role in the system are more favourable to E-LIS coverage (Editors $\bar{x} = 3.88$, $SD = 1.11$ and Registered Users $\bar{x} = 3.75$, $SD = 0.94$), while Unregistered Users are not completely satisfied by the content quantity found in it ($\bar{x} = 3.09$, $SD = 1.09$). Those who reported very rare use disagreed with the statement of satisfaction of coverage ($\bar{x} = 2.67$, $SD = 1.17$), while those who reported very often use stated also their agreement with the statement ($\bar{x} = 4.12$, $SD = 1.05$). ANOVA showed that the role and usage factors differentiate significantly the behaviour of the sub-samples ($p < .001$ for all features in the usefulness category).

5.2.2. Usability questions

In the usability category (2.1–2.5), Ease of Use and Learnability scored very high mean rates (4.06 and 4.08, respectively). Once again users with active presence in the archive consider E-LIS as very easy to use (Editors $\bar{x} = 4.00$, $SD = 0.79$ and Registered Users $\bar{x} = 4.41$, $SD = 0.76$), but those with little interference with formal roles do not share the same opinion to the same extend (Unregistered Users $\bar{x} = 3.86$, $SD = 1.18$). Similar were the results in the Learnability attribute, where Editors ($\bar{x} = 4.12$, $SD = 0.86$) and Registered Users ($\bar{x} = 4.36$, $SD = 0.72$) believe that E-LIS is a fairly learnable system, while the opinion of the Unregistered Users was the same, but to a lesser degree ($\bar{x} = 3.89$, $SD = 1.10$). Participants reporting very rare use do not consider E-LIS as an easy to use DL ($\bar{x} = 3.17$, $SD = 1.27$) and they were in contrast with those who reported very often use ($\bar{x} = 4.44$, $SD = 0.65$). Participants with very often use reported that E-LIS is an easy to learn DL as well ($\bar{x} = 4.32$, $SD = 0.75$), but the participants with very rare use argued on its learnability ($\bar{x} = 3.08$, $SD = 1.28$).

A non-significant difference between the sub-groups representing different roles was traced in the case of Learnability ($F = 2.314$, $p > .05$), while all other features showed significant differences at $p < .001$ for Aesthetic ($F = 7.410$) and Terminology ($F = 5.558$) and $p < .05$ for Ease of Use ($F = 2.781$) and Navigation ($F = 3.023$). Significant differences were also found between the sub-groups based on usage at the levels of $p < .001$ for Ease of Use ($F = 8.906$), Aesthetic ($F = 7.806$) and Navigation ($F = 5.820$) and $p < .05$ for Terminology ($F = 4.619$) and Learnability ($F = 3.646$).

5.2.3. Performance questions

In the Performance category (3.1–3.3), the two dominating IR evaluation measures, Precision and Recall, did not capture the preference of the participants. They did not supply a clear answer for their satisfaction with the Precision and the Recall of E-LIS. While Editors' responses for Precision were above the mean value ($\bar{x} = 3.65$, $SD = 1.22$) and were in accordance with the rates provided by Registered Users ($\bar{x} = 3.61$, $SD = 0.92$), Unregistered Users' rates were below medium showing a disagreement about the ability of E-LIS to return precise results ($\bar{x} = 2.85$, $SD = 0.98$). Similar were the results for Recall. Editors ($\bar{x} = 3.65$, $SD = 1.00$) and Registered Users ($\bar{x} = 3.48$, $SD = 0.88$) agreed that E-LIS could provide them the right number of results, but Unregistered Users demonstrated a negative attitude towards this feature ($\bar{x} = 2.94$, $SD = 0.96$). Participants that reported very rare use believed that both Precision ($\bar{x} = 2.42$, $SD = 1.18$) and Recall ($\bar{x} = 2.46$, $SD = 0.98$) are not satisfactory. The participants with more frequent use of E-LIS declared their mediocre satisfaction with Precision ($\bar{x} = 3.76$, $SD = 1.05$) and Recall ($\bar{x} = 3.84$, $SD = 0.90$).

Significant differences were indicated at the level of $p < .001$ for Precision ($F = 6.521$) and Response Time ($F = 5.501$), and $p < .05$ for Recall ($F = 4.446$). Participants with very low frequency of usage showed their disagreement with the statements of satisfaction in Precision ($\bar{x} = 2.42$, $SD = 1.18$) and Recall ($\bar{x} = 2.46$, $SD = 0.98$). On contrast participants with high levels of usage agreed with the respective statements and provided scores above medium ($\bar{x} = 3.76$, $SD = 1.05$ for Precision and $\bar{x} = 3.84$, $SD = 0.90$ for Recall). Significant differences were reported for all types of reported usage ($p < .001$).

5.3. Inter-axes correlations

Although the presented results reflect participants’ opinions about the DLs features, the correlations that they assign to them and their categories were also investigated. Table 5 presents the inter-attribute Pearson correlations, all of which were found to be significant, ranging from $r = 0.35$ to $r = 0.79$. An overview of the table demonstrates significant correlations between Reliability and Format, Reliability and Level, and Coverage and Level for the usefulness category, Ease of Use and Navigation, Ease of Use and Learnability, Navigation and Aesthetics and Terminology and Learnability for the usability category and Recall and Precision for the performance category.

Similarly, Fig. 2 presents the inter-category correlation values. Participants correlate highly all categories between each other (all correlations are above 0.70), with most notable the correlation between usefulness

Table 5
Inter-item correlation matrix

	1.1	1.2	1.3	1.4	1.5	2.1	2.2	2.3	2.4	2.5	3.1	3.2	3.3
1.1	1												
1.2	0.62	1											
1.3	0.63	0.74	1										
1.4	0.57	0.65	0.72	1									
1.5	0.44	0.50	0.61	0.65	1								
2.1	0.58	0.63	0.61	0.57	0.49	1							
2.2	0.51	0.61	0.52	0.48	0.44	0.74	1						
2.3	0.54	0.61	0.59	0.52	0.37	0.79	0.76	1					
2.4	0.58	0.67	0.59	0.64	0.47	0.70	0.65	0.71	1				
2.5	0.54	0.62	0.57	0.50	0.47	0.76	0.63	0.73	0.78	1			
3.1	0.56	0.53	0.67	0.63	0.59	0.54	0.51	0.54	0.54	0.50	1		
3.2	0.52	0.67	0.58	0.58	0.48	0.61	0.54	0.55	0.66	0.66	0.52	1	
3.3	0.57	0.40	0.53	0.50	0.56	0.45	0.35	0.38	0.40	0.45	0.58	0.49	1

$p < 0.01$

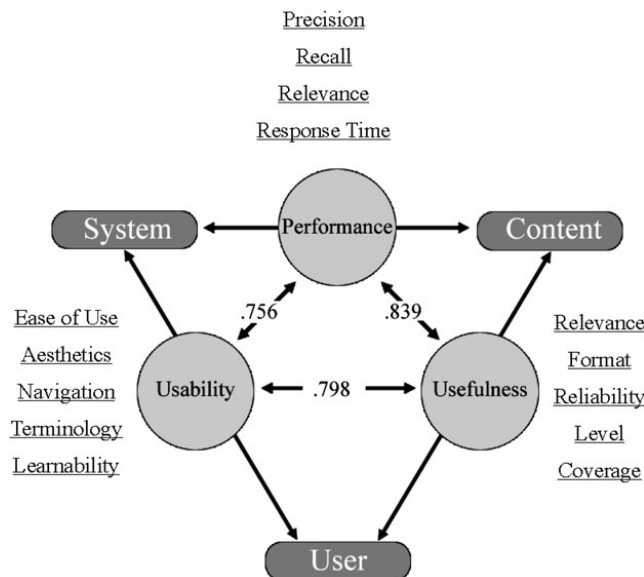


Fig. 2. Correlations between the evaluation axes.

and performance ($r = .839$). In regard to the relation between usefulness and usability, the high correlation ($r = .798$) verifies the assumption that the quality of user interaction depends strongly on both axes.

5.4. Predictive strength of factors

Stepwise multiple regression analysis was performed for each category of constructs in order to estimate their predictive strength. The attributes of each category were entered as independent variables and the regression model was performed on a dependent variable that measured the overall quality of each category. Table 6 presents the results of stepwise multiple regressions for each category. The multicollinearity values of VIF and Tolerance were calculated for each variable of the regression model. This check showed that all values were below 10 (min. = 1.496, max. = 3.369) and Tolerance were found more than .2, hence data multicollinearity was evaded.

Relevance ($t(125) = 4.697, p < .001$), Level ($t(125) = 3.956, p < .001$) and Coverage ($t(125) = 2.887, p < .01$) had a significant effect on the prediction for DL usefulness. The three attributes are responsible for the 65% of the observed variance. Four out of five usability attributes of the usability category produce the 63% of variance and have a significant effect at the $p < .01$ level, while the fifth construct, Navigation, was excluded from the analysis ($t(126) = .534, p > .05$). The data in the Performance category reveal that Precision ($t(127) = 8.050, p < .001$) and Recall ($t(127) = 4.467, p < .001$) cause 66% of variance, when the remaining construct, Response Time ($t(127) = 1.713, p > .05$), has a non-significant effect.

5.5. E-LIS functionalities evaluation

Table 7 presents the participants' opinion on the questions of the third part of the questionnaire regarding E-LIS functionalities and characteristics and how they affect the system's overall usefulness and usability. The selected functionalities reflect adequately the E-LIS nature and its information management processes, which are the information discovery methods (e.g. search and browse), personalized delivery of services, peripheral services provision (e.g. e-mail alerts, cross-reference), its OA nature and finally the procedural phases (e.g. submission, review, edit, delete of contributions, etc.). Participants exhibited a positive attitude towards all these functionalities (mean ranges from 3.31 to 4.08). In particular 50% of the participants strongly believe that E-LIS is both useful and usable to them due to its open access (OA) nature, but they also believe in their majority that the personalized functionalities of E-LIS are not quite useful and usable.

Table 6
Multiple regression analysis

	<i>B</i>	SE <i>B</i>	β	<i>p</i>
<i>Usefulness</i>				
$R^2 = 0.653$, adjusted $R^2 = 0.639$				
Relevance	0.346	0.074	0.338	<.001
Format	0.053	0.090	0.049	>.05
Reliability	0.034	0.102	0.031	>.05
Level	0.342	0.086	0.340	<.001
Coverage	0.220	0.076	0.208	<.01
<i>Usability</i>				
$R^2 = 0.634$, adjusted $R^2 = 0.622$				
Ease of use	0.245	0.091	0.267	<.01
Aesthetic	0.159	0.077	0.173	<.01
Terminology	0.187	0.087	0.198	<.01
Learnability	0.248	0.096	0.256	<.01
<i>Performance</i>				
$R^2 = 0.665$, adjusted $R^2 = 0.657$				
Precision	0.494	0.061	0.539	<.001
Response time	0.104	0.061	0.108	>.05
Recall	0.292	0.065	0.294	<.001

Table 7
Descriptive statistics for functionalities questions (1 = disagree. 5 = agree)

	1 n (%)	2 n (%)	3 n (%)	4 n (%)	5 n (%)	\bar{x}	SD
Browsing usefulness	5 (3.8)	9 (6.9)	24 (18.3)	52 (39.7)	41 (31.3)	3.88	1.053
Browsing usability	5 (3.8)	8 (6.1)	35 (26.7)	43 (32.8)	40 (30.5)	3.80	1.063
Search usefulness	5 (3.8)	7 (5.3)	33 (25.2)	46 (35.1)	40 (30.5)	3.83	1.046
Search usability	5 (3.8)	8 (6.1)	37 (28.2)	41 (31.3)	40 (30.5)	3.79	1.067
Personal account usefulness	19 (14.5)	16 (12.2)	33 (25.2)	31 (23.7)	32 (24.4)	3.31	1.354
Personal account usability	18 (13.7)	18 (13.7)	31 (23.7)	34 (26.0)	30 (22.9)	3.31	1.335
Services usefulness	15 (11.5)	5 (3.8)	31 (23.7)	42 (32.1)	38 (29.0)	3.63	1.260
Services usability	15 (11.5)	6 (4.6)	32 (24.4)	45 (34.4)	33 (25.2)	3.57	1.241
OA usefulness	9 (6.9)	3 (2.3)	23 (17.6)	29 (22.1)	67 (51.1)	4.08	1.183
OA usability	9 (6.9)	5 (3.8)	26 (19.8)	30 (22.9)	61 (46.6)	3.98	1.202
Procedures usefulness	10 (7.6)	7 (5.3)	36 (27.5)	44 (33.6)	34 (26.0)	3.65	1.150
Procedures usability	9 (6.9)	7 (5.3)	41 (31.3)	38 (29.0)	36 (27.5)	3.65	1.143

Stepwise multiple regression analysis was repeated including as independent variables the system functionalities in investigation, to predict the overall usefulness and usability. Despite the fact that five of six functionalities (browsing, search, personal account, services and OA) are considered to cause the 55% of variance ($R^2 = 0.556$, $F = 31.312$, $p < .001$), only two of the functionalities possess significant strength to predict satisfaction on usefulness. In particular participants believe that the services provided by E-LIS ($t(125) = 3.227$, $p < .05$) and the fact that it is an OA system ($t(125) = 4.152$, $p < .01$) are enough to predict usefulness.

The OA nature of E-LIS was strengthened even more as it was conceived by the participants as a feature that could predict the overall usability ($t(125) = 2.878$, $p < .01$), together with the personalized functionalities that the e-print archive provides ($t(125) = 2.317$, $p < .05$). Once again the above mentioned five functionalities are considered to capture 57% of usability variance ($R^2 = 0.579$, $F = 34.331$, $p < .001$).

5.6. Synopsis of results

In summary the results show that usefulness and usability are two related concepts that jointly affect user satisfaction. However the gravity of system performance was also exhibited and demanded adequate attention even in user-centered evaluations. Participants of the current study believe that usefulness can be predicted by the Open Access to relevant, levelled and widely covered information supported as well by peripheral services. They also believe that E-LIS is considered as a usable system if it satisfies the conditions of Ease of Use, Aesthetics, Terminology and Learnability. According to the participants, the level of usability is increased by the fact that E-LIS is an Open Access system and that provides certain personalized features.

6. Discussion

The present study reported differences in users' opinions as their experience changes. Users' experience, reflected in the reported usage and the role they have in the system, affects their perceptions about E-LIS usefulness and usability. This conclusion validates prior studies' findings (Koohang & Ondracek, 2005) and makes them applicable in the field of digital repositories. Differences between experienced and novice users were also reported for performance of repositories usage (Silva, Laender, & Gonçalves, 2005). Whilst one can support that there are studies claiming no actual differences in the user performance between novice and experienced users of DLs (Kengeri, Seals, Harley, Reddy, & Fox, 1999), this does not infer a clear conflict among the various experimental findings, but indicates the need to exploit a wide range of methods and metrics, especially in the usability field. The inclusion of editors and the administrative team aimed to strengthen the concept of experience and to provide a more detailed analysis, since the ANOVA test showed significance differences ($F = 9.016$, $p > .01$). Editors represent an important part of the mechanical process of self-archiving and possess knowledge of the system and content parameters. Despite argumentation on the bias that they may cause, the opinion of representative classes should be gathered as they typify different needs, requirements and behaviours. Therefore, the favourable attitude they expressed towards E-LIS' usefulness and usability

seems logical and explainable at the same time. In general, scholars from this discipline will apply some sort of bias on research studies, because of the high level of awareness on self-archiving processes. Swan and Brown (2005) have reported that 60% of the sample from this discipline, which was the highest among the various disciplines, knows what self-archiving encapsulates and that is a means towards OA publishing.

Concerning the other users' characteristics, participants appoint mediocre importance to E-LIS when they execute their information seeking tasks. This low level of significance showed respectable differences only for the factor of usage ($F = 40.399$, $p = 0.001$), which leads to the conclusion that system usage may affect the participants' preference. Participants who frequently use E-LIS consider it as a regular channel for retrieving information and quite naturally assign to it high priority. Furthermore, users are committed to spend as much time and effort is needed in using E-LIS to find the information they want. Once again, users' behaviour was significantly different between these two parameters (for Time $F = 96.456$, $p = 0.001$, while for Effort $F = 95.400$, $p = 0.001$). Finally the introduction channels suggest an elaboration of the promotional and marketing means that the administration team employs. These results, which are in agreement with the results from other studies (Allen, 2005; Hitchcock et al., 2002), request much more attention in the area of support and documentation, especially in cases of subject-oriented repositories, where spatial distribution restricts physical interaction with the served community.

Most of the usefulness attributes were highly appreciated by the users. However participants are not satisfied by E-LIS Coverage. Apparently they prefer more uploaded content, ranging to many years, and their preference is in accordance to the remarks of the Citebase subjects (Hitchcock et al., 2002). As the authors of the Citebase research note, the responsibility for content growth is transferred to the hands of the users themselves. In addition raw statistic evidence, provided by the E-LIS administrative team, showed a steady increase in the collection size. Shearer (2003) also highlights the indissoluble connection between input activity and usage. This vice-versa relation explains the entrustment of depositing works in a repository of high usage and wide recognition, as well as the rate of usage according to the deposited material.

Users seem to prefer systems that provide them a structured and levelled presentation of information, which is relevant to their information tasks. However they do not appoint significance to the role of Format and Reliability. In the former attribute, this interesting exclusion can be credited to the deposit policy of the DL, which allows only the PDF file format. The result for the latter attribute can be explained by the fact that the only prerequisite for the documents deposit in E-LIS is relevance and document integrity. Therefore it can be supported that users of such systems do not assign significance to reliability, due to their awareness that the content review and approval procedure does not employ reliability criteria. It is remarkable that both attributes are highly correlated ($r = 0.74$).

The five usability features scored high above the medium. Users reported their preference to Easiness of Use and Learnability, while understandable Terminology and Aesthetics do not attain high levels of users' preference. This allocation of preference is in accordance with conclusions by other studies, such as Kani-Zabihi et al. (2006). According to their findings, users give higher priority to the easy discovery of information and the high degree of familiarization with DL functionalities. Moreover, learnability remains one of the key factors for the acceptance of these systems and that is linked with the prospective benefits and the added value of the application (Davis & Connolly, 2007), while in the present study ANOVA demonstrates that learnability is the only widely preferred attribute, regardless the participants' roles. In spite Navigation gained high scores by participants, it was not perceived as an important feature that affects overall opinion on DL usability. However literature reveals that navigation is considered an important factor that influences many of the traditional work tasks of librarians and information scientists (Moyo, 2002). One possible explanation might be that the information scientists with the assistance of appropriate navigation tools and aids (such as indices, menu bars, etc.) have gained the experience needed and may overtake any obstacles in their navigation routes. This experience is underlined by some DL development studies that highlight the role of experienced users, in the successful information discovery through navigation tools (Xu, 2004).

Although not a primary aim of this study, following the theoretic foundations of ITF, the performance axis was also analyzed. Precision and Recall are considered the main factors that affect users' opinion on system performance. Despite the dependence on factors that influence Response Time (e.g. traffic, webpage size, etc.) and the critical position in the evaluation of web-based systems, users expressed higher levels of preference to the "traditional" measurements of Precision and Recall.

It is very intriguing the fact that participants in this study believe that the open access nature of e-print archives makes them both useful and usable. This may be associated with the current trends of the scientific community to advocate open access. Participants in this study did not appreciate the provision of personalized services, but they were satisfied with the usefulness and usability of the browsing and searching features. These features allow users to retrieve documents in a wide number of methods, such as author, country or year indices and advanced search interfaces. In the comparative usability study by Kim (2005) subjects using Eprints, upon which E-LIS is build, were more satisfied with the retrieval functionalities and made less time to conclude their tasks, especially in tasks employing browsing and searching in simple search interfaces. Therefore, it can be concluded that the wealth and structure of retrieval tools outweighs the inefficiency of Eprints software to search on a full-text level, as reported by Goh et al. (2006).

Furthermore participants in this study correlated very highly the Usefulness and Performance categories. A possible explanation is that participants are correlating particularly usefulness and performance due to their knowledge of the importance of these two concepts in the effective operation of a DL (also demonstrated in Tsakonias & Papatheodorou, 2006). As being information scientists that manage information objects and handle them under different criteria that are equivalent with the usefulness attributes of ITF (e.g. relevance and format), they know how these support users' work tasks. At the same time they are aware of the significance of IR evaluation criteria for the successful completion of their information searching activities. This high correlation value demonstrates that, despite the focus of the current study, user satisfaction is also dependent on the classic IR criteria and that content attributes are affecting system performance.

7. Conclusions

While commercial DLs have proved their degree of self-sustainability, e-print archives' operation is dependent on many issues, either political, or economic. One of the major challenges that e-prints face is to become self-sustainable systems closely linked with users' work tasks, instead of gradually transform into graveyards of invaluable documents. Although the issue of open access systems has gained significant attention and the consensus for their evaluation strengthens among the scientific community, there are not many studies concentrating on usefulness and usability issues. In this study a theoretical model for DL evaluation was applied to assess the usefulness and usability of an OA digital library and it was attempted to demonstrate the significant content and system attributes and how they affect user interaction and satisfaction. These attributes request research attention and in depth analysis on each one will reveal more details and contribute to the overall evaluation of DLs.

Acknowledgements

The authors wish to thank Antonella DeRobbio, Imma Subirats, Zeno Tajoli, E-LIS administrators, and the national editors of E-LIS for their invaluable help and support during the realization of this study, Dimitris Gavrilis, Department of Electrical Engineering, University of Patras, Patras, Greece, for his technical assistance, and Tasos Tsagkos for his constructive comments on the statistic analysis of this study.

References

- Allen, M. (2002). A case study of the usability testing of the University of South Florida's virtual library interface design. *Online Information Review*, 26(1), 40–53.
- Allen, J. (2005). Interdisciplinary differences in attitudes towards deposit in institutional repositories. Master Thesis in Department of Information and Communications, Manchester Metropolitan University (UK). <http://eprints.rclis.org/archive/00005180/> (retrieved 27.04.07).
- Andrew, T. (2003). Trends in self-posting of research material online by academic staff. *Ariadne*, 37, <http://www.ariadne.ac.uk/issue37/andrew/intro.html> (retrieved 27.04.07).
- Bertot, J. C., Snead, J. T., Jaeger, P. T., & McClure, C. R. (2006). Functionality, usability, and accessibility: Iterative user-centered evaluation strategies for digital libraries. *Performance Measurement and Metrics*, 7(1), 17–28.
- Brennan, P. B. M., Burkhardt, J., McMullen, S., & Wallace, M. (1999). What does electronic full-text really mean? A comparison of database vendors and what they deliver. *Reference Services Review*, 27(2), 113–126.
- Center for Research Libraries (2006). General factors to consider in evaluating digital repositories. *Focus*, 25(2), 3–4.

- Chowdhury, S., Landoni, M., & Gibb, F. (2006). Usability and impact of digital libraries: A review. *Online Information Review*, 30(6), 656–680.
- Cockrell, B. J., & Jayne, E. A. (2002). How do I find an article? Insights from a web usability study. *Journal of Academic Librarianship*, 28(3), 122–132.
- Correia, A. M. R., & Castro Neto, M. (2002). The role of eprint archives in the access to, and dissemination of, scientific grey literature: LIZA—a case study by the National Library of Portugal. *Journal of Information Science*, 6(28), 231–241.
- Covey, D. T. (2002). *Usage and usability assessment: Library practices and concerns*. Washington, DC: Digital Library Federation, CLIR.
- Davis, P. M., & Connolly, M. J. L. (2007). Institutional repositories: Evaluating the reasons for nonuse of Cornell University's installation of DSpace. *D-Lib Magazine*, 13(3/4), <http://dlib.org/dlib/march07/davis/03davis.html> (retrieved 29.05.07).
- Delone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the depending variable. *Information Systems Research*, 3(1), 60–95.
- Ebenezer, C. (2003). Usability evaluation of an NHS library website. *Health Information and Libraries Journal*, 20(3), 134–142.
- Ferreira, S. M., & Pithan, D. N. (2005). Usability of digital libraries: A study based on the areas of information science and human–computer interaction. *OCLC Systems and Services*, 21(4), 311–323.
- Fuhr, N., Hansen, P., Mabe, M., Micsik, A., & Sølvberg, I. (2002). Digital libraries: A generic classification and evaluation scheme. *Proceedings of the 5th European conference on Digital Libraries 2001: LNCS* (Vol. 2163, pp. 187–199). Berlin: Springer.
- Fuhr, N., Tsakonias, G., Aalberg, T., Agosti, M., Hansen, P., Kapidakis, S., et al. (in press). Evaluation of digital libraries. *International Journal on Digital Libraries*, doi:10.1007/s00799-007-0011-z.
- Fuller, D. M., & Hinegardner, P. G. (2001). Ensuring quality website redesign: The University of Maryland's experience. *Bulletin of the Medical Library Association*, 89(4), 339–345.
- Goh, D. H., Chua, A., Khoo, D. A., Khoo, E. B., Mak, E. B., & Ng, M. P. (2006). A checklist for evaluating open source digital library software. *Online Information Review*, 30(4), 360–379.
- Hartson, R. H., Shivakumar, P., & Pérez-Quinones, M. A. (2004). Usability inspection of digital libraries: A case study. *International Journal on Digital Libraries*, 4(2), 108–123.
- Heery, R., & Anderson, S. (2005). Digital repositories review. Report by UKOLN and AHDS. http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf (retrieved 24.01.07).
- Hitchcock, S., Woukeu, A., Brody, T., Carr, L., Hall, W., & Harnad, S. (2002). Evaluating Citebase, an open access Web-based citation-ranked search and impact discovery service. <http://opcit.eprints.org/evaluation/Citebase-evaluation/evaluation-report.html> (retrieved 24.01.07).
- Hong, W., Thong, J. Y. L., Hong, W., & Tam, K. (2002). Determinants of user acceptance of digital libraries: an empirical examination of individual differences and system characteristics. *Journal of Management Information Systems*, 18(3), 97–124.
- Jackson, M. (2001). A user centered approach to the evaluation of a hybrid library. *Performance Measurement and Metrics*, 2(2), 97–107.
- Järvelin, K., & Ingwersen, P. (2004). Information seeking research needs extension towards tasks and technology. *Information Research*, 10(1), <http://InformationR.net/ir/10-1/paper212.html> (retrieved 24.01.07).
- Jeng, J. (2005). What is usability in the context of digital library and how it can be measured?. *Information Technology and Libraries* 24(2), 47–56.
- Kani-Zabihi, E., Ghinea, G., & Chen, S. Y. (2006). Digital libraries: What do users want? *Online Information Review*, 30(4), 395–412.
- Kengeri, R., Seals, C. D., Harley, H. D., Reddy, H. P., & Fox, E. A. (1999). Usability study of digital libraries: ACM, IEEE-CS, NCSTRL, NDLTD. *International Journal on Digital Libraries*, 2(2–3), 157–169.
- Kim, K. (2002). A model-based approach to usability evaluation for digital libraries. In A. Blandford & G. Buchanan (Eds.) *JCDL'02 Workshop on usability of Digital Libraries*. <http://www.ucl.ac.uk/annb/docs/Kim33.pdf> (retrieved 27.04.07).
- Kim, J. (2005). Finding documents in a digital institutional repository: DSpace and Eprints. *Proceedings of the American Society for Information Science and Technology*, 42(1).
- Kim, J. (2006). Motivating and impeding factors affecting faculty contribution to institutional repositories. In *Digital curation and trusted repositories: Seeking success workshop, June 15, 2006 in conjunction with Joint conference on Digital Libraries (JCDL 2006, June 11–15, 2006 – Chapel Hill, NC, USA)*. <http://sils.unc.edu/events/2006jcdl/digitalcuration/Kim-JCDLWorkshop2006.pdf> (retrieved 29.05.07).
- Kobayashi, M., & Takeda, K. (2000). Information retrieval on the web. *ACM Computing Surveys*, 32(2), 144–173.
- Koohang, A., & Ondracek, J. (2005). Users' views about the usability of digital libraries. *British Journal of Educational Technology*, 36(3), 407–423.
- Krottmaier, H. (2002). Current and future features of digital journals. In M. H. Shafazand & A. M. Tjoa (Eds.), *EurAsia-ICT 2002: Information and communication technology: First EurAsian Conference: LNCS. Vol. 2510* (pp. 495–502). Berlin: Springer.
- Kurata, K., Matsubayashi, M., Mine, S., Muranushi, T., & Ueda, S. (2006). Electronic journals and their unbundled functions in scholarly communication: Views and utilization by scientific, technological and medical researchers in Japan. *Information Processing and Management*, 43(5), 1402–1415.
- Kyrillidou, M., & Giersch, S. (2005). Developing the DigiQUAL protocol for digital library evaluation. In *Proceedings of the 5th ACM/IEEE-CS joint conference on Digital Libraries* (pp. 172–173). New York, NY: ACM Press.
- Landoni, M., & Bell, S. (2000). Information retrieval techniques for evaluating search engines: A critical overview. *Aslib Proceedings*, 52(3), 124–129.
- Liaw, S., & Huang, H. (2003). An investigation of user attitudes towards search engines as an information retrieval tool. *Computers in Human Behavior*, 19(6), 751–762.
- Liu, Z. (2004). Perceptions of credibility of scholarly information on the web. *Information Processing and Management*, 40(6), 1027–1038.

- Luce (2001). E-prints intersect the digital library: Inside the Los Alamos arXiv. *Issues in Science and Technology Librarianship* (29). <http://www.istl.org/01-winter/article3.html> (retrieved 29.05.07).
- McMullen, S. (2001). Usability testing in a library Web site redesign project. *Reference Services Review*, 29(1), 7–22.
- Mercer, L. S. (2000). Measuring the use and the value of electronic journals and books. *Issues in science and technology librarianship*, 25, <<http://www.istl.org/00-winter/article1.html>> (retrieved 11.08.07).
- Moyo, L. M. (2002). Collections on the Web: Some access and navigation issues. *Library Collections, Acquisitions, and Technical Services*, 26(1), 47–59.
- Nicholas, D., Huntington, P., & Rowlands, I. (2005). Open access journal publishing: the views of some of the world's senior authors. *Journal of Documentation*, 61(4), 497–519.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Park, S. (2000). Usability, user preferences, effectiveness, and user behaviors when searching individual and integrated full-text databases: Implications for digital libraries. *Journal of the American Society for Information Science*, 51(5), 456–468.
- Rowlands, I., & Nicholas, D. (2005). Scholarly communication in the digital environment: The 2005 survey of journal author behaviour and attitudes. *Aslib Proceedings: New Information Perspectives*, 57(6), 481–497.
- Saracevic, T. (2004). Evaluation of digital libraries: An overview. In *Notes of the DELOS WP7 workshop on the evaluation of Digital Libraries, Padua, Italy*. http://dlib.ionio.gr/wp7/workshop2004_program.html (retrieved 30.04.07).
- Shearer, K. (2003). Institutional repositories: Towards the identification of critical success factors. *Canadian Journal of Information and Library Science*, 27(3), <http://hdl.handle.net/1880/43357> (retrieved 02.05.07).
- Silva, L. V., Laender, A. H. F., & Gonçalves, M. A. (2005). A usability evaluation study of a digital library self-archiving service. In *Proceedings of the 5th ACM/IEEE-CS joint conference on digital libraries* (pp. 176–177). New York, NY: Denver, CO, USA, ACM Press.
- Silva, L. V., Laender, A. H. F., & Gonçalves, M. A. (2007). Evaluating a digital library self-archiving service: The BDBComp user case study. *Information Processing and Management*, 43(4), 1103–1120.
- Stelmaszewska, H., & Blandford, A. (2002). Patterns of interactions: User behaviour in response to search results. In A. Blandford & G. Buchanan (Eds.), *JCDL '02 Workshop on usability of digital libraries*. <<http://www.ucl.ac.uk/annb/docs/Stelmaszewska29.pdf>> (retrieved 11.08.07).
- Sutcliffe, G., Ennis, M., & Hu, J. (2000). Evaluating the effectiveness of visual user interfaces for information retrieval. *International Journal of Human-Computer Studies*, 53(5), 741–763.
- Swan, A., & Brown, S. (2005). *Open access self-archiving: An author study*. Key Perspectives Limited.
- Theng, Y. L., Duncker, E., Mohd-Nasir, N., Buchanan, G., & Thimbleby, H. (1999). Design guidelines and user-centred digital libraries. In *Proceedings of the 3rd European conference on Digital Libraries 2001: LNCS. Vol. 1696* (pp. 167–183). Berlin: Springer.
- Thong, J. Y. L., Hong, W., & Tam, K. (2002). Understanding user acceptance of digital libraries: What are the roles of interface characteristics, organizational context, and individual differences? *International Journal of Human-Computer Studies*, 57(3), 215–242.
- Toms, E. G. (2002). Information interaction: Providing a framework for information architecture. *Journal of the American Society for Information Science and Technology*, 53(10), 855–862.
- Tsakonias, G., Kapidakis, S., & Papatheodorou, C. (2004). Evaluation of user interaction in digital libraries. In *Notes of the DELOS WP7 workshop on the evaluation of Digital Libraries, Padua, Italy*. http://dlib.ionio.gr/wp7/workshop2004_program.html (retrieved 24.01.07).
- Tsakonias, G., & Papatheodorou, C. (2006). Analysing and evaluating usefulness and usability in electronic information services. *Journal of Information Science*, 32(5), 400–419.
- Vakkari, P., & Hakala, P. (2000). Changes in relevance criteria and problem stages in task performance. *Journal of Documentation*, 56(5), 540–562.
- Van House, N. A., Butler, M. H., Ogle, V., & Schiff, L. (1996). User-centered iterative design for digital libraries: The Cypress experience. *D-Lib Magazine*, 2(2), <http://www.dlib.org/dlib/february96/02vanhouse.html> (retrieved 24.01.07).
- Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation*, 55(3), 249–270.
- Wolfram, D., & Xie, H. (2002). Traditional IR for web users: A context for general audience digital libraries. *Information Processing and Management*, 38(5), 627–648.
- Wyles, R., Maxwell, M., & Yamog, J. (2006). Technical evaluation of selected open source repository solutions. <https://eduforge.org/docman/view.php/131/1062/RepositoryEvaluationDocument.pdf> (retrieved 24.01.07).
- Xie, H. (2004). Online IR system evaluation: Online databases versus Web search engines. *Online Information Review*, 28(3), 211–219.
- Xie, H. (2006). Evaluation of digital libraries: Criteria and problems from users' perspectives. *Library and Information Science Research*, 28(3), 433–452.
- Xu, Q. (2004). Content management and resources integration: a practice in Shanghai Digital Library. In *Proceedings the 7th international conference on Asian Digital Libraries, ICADL 2004, Digital Libraries: International collaboration and cross-fertilization: LNCS. Vol. 3334* (pp. 25–34). Berlin: Springer.
- Yang, Z., Cai, S., Zhou, Z., & Zhou, N. (2004). Development and validation of an instrument to measure user perceived service quality of information presenting web portals. *Information and Management*, 42(4), 575–589.

Giannis Tsakonias holds a BSc from the Department of Archive and Library Sciences, Ionian University, Corfu, Greece and he is a PhD candidate in the same department. He also is librarian in the Library and Information Service, University of Patras, Greece. His research interests include user centered digital library evaluation, information behaviour and visual communication.

Christos Papatheodorou holds a BSc and a PhD in Computer Science from the Department of Informatics, Athens University of Economics and Business, Athens, Greece. He is an Assistant Professor at the Department of Archive and Library Sciences, Ionian University, Corfu, Greece, where he teaches Information Systems, Metadata and Information Retrieval. His research interests include, user-centered digital library evaluation, user modelling, web mining and metadata interoperability.