

Papatheodorou C. & Tsakonas, G. (2007). Exploring Usefulness and Usability In the Evaluation of Open Access Digital Libraries. *Information Processing and Management*, 44 (2008). 1234-1250. Retrieved March 24, 2008.

## Summary

In 2007, Greek researchers Christos Papatheodorou and Giannis Tsakonas published their study about open access (OA) digital libraries (DL). Instead of focusing on the publishing or economic perspectives of user-centered design of DLs, their study focuses on OA publishing and how the format fulfills user needs. This qualitative study explores the most significant content and system features affect the usefulness and usability of the OA e-print archive, E-LIS. This international, free access online archive holds full-text documents and metadata commonly used by researchers of library and information sciences. It also provides for document submission within user roles and rights, and is edited by one or more people in each participating country.

By using a previously developed theoretical DL evaluation method, the researchers explored the important attributes and their functionality in the OA E-LIS. By comparing three main components of DL interaction – the system, content, and user – an axis and relationship developed to describe the components' interactions. The user base is comprised of registered and non-registered users. Accomplished through the use of a 34 question online questionnaire, 131 individual users of the E-LIS site provided the interaction results for this study. The questions explored usefulness, usability, and performance, providing a way for the researchers to explore their inter-axes correlations.

The results of this study showed that there were almost equal parts registered and unregistered E-LIS users with the majority of users residing in Europe. The usage rates had a normal distribution, and site usage increased with the role, with registered users at the top and unregistered users at the bottom. E-LIS users showed that the format was the leader of the Usefulness category, ease of use was highest in the Usability category, and response time ranked highest in Performance. Of these three categories, usefulness rated highest while Performance rated lowest. All the functionalities of the e-print archive rated high among all users, with most rating the OA usefulness and usability the highest of the twelve presented in the questionnaire.

From this research, Papatheodorou and Tsakonas showed that both usefulness and usability affect user satisfaction. It also gave credibility to OA usefulness if casual users support the information, and the system satisfies the conditions of ease of use, aesthetics, terminology, and learnability. User-personalized features can increase this. Users' opinions changed as their role and experience in the online community changes. Usefulness and Performance categories were also shown to possess high correlation. All of these factors

contribute to the e-print archives self-sustainability and showed that the theoretical DL evaluation method could be used to interpret and explore OA DLs.

### **Review of Previous Research**

The exhaustive literature review consisted of mostly recent research. Since most of the studies historically reside in the commercial and privatized sectors of DLs, the researchers acknowledged this fact early in the review. The research consists of credible primary sources and research, consistently cited in the article and expanded upon in the References section. Limited by the recent advent of their study subject, the researchers looked for studies that related to individual areas of the E-LIS repository. The literature review covers three main areas, OA systems, DL evaluation methods, and the interaction between information behavior and human-computer interactions.

First, an investigation was conducted in the interaction between information behavior and human-computer interactions. Papatheodorou and Tsakonas gave examples of relationships between the user and information resources' structural design and its effect of supporting or preventing usability. They also provided previous studies of interaction between the content and system with a tri-polar structure, much like the one created for the study. The referenced study focused on the design of DLs and how it affects information searching strategies. Another focused on usefulness and usability as viewed in the opinions of users in the five categories of usability, collection quality, service quality, system performance efficiency, and user opinions. That methodology provided a base for a section the survey developed for this research.

Second, the researchers explored DL evaluation models. Papatheodorou and Tsakonas discussed a number of methods, the Technology Acceptance Model, DigiQUAL, and two other un-named methods proposed by other leading researchers. Each of the methods provided different parameters and areas of evaluation, and all try to find common ground in which to study DL systems. All them contained similar components as the theoretical groundwork for the Interaction Triptych Framework (ITF) model, developed in 2004 by the authors and their colleagues.

Third, researchers cited previous OA systems research. They were quick to acknowledge that there are few studies regarding users' interactions with these systems. The research showed support for ideas of the OA subject-based e-print server system an attractive tool for physicists, computer scientists, economist, as well as other scholars. These scholars favored the self-archiving tool and the idea of an alternative to traditional publishing means. Electronic journals and institutional repositories showed most of OA interest, and suggested a strong link with the traditional structures of scholars, libraries, funding agencies, and both government and non-governmental organizations. This supported the researcher's selection of study participants from all types of E-LIS users.

## **Problem and Purpose**

Papatheodorou and Tsakonas seek to explore and define the relationships and factors between usefulness, useability, and performance in an OA e-print archive. The researchers clearly present the research problem, "... 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" (Tsakonas, 2007). This is specifically stated in the first paragraph of section 3. The purpose is somewhat tied to the literature review through the explorations of evaluation methods, OA research, and interactions between information behavior and human-computer interactions. Using the rationale that this study is exploring a new area of information archiving and e-prints, this literature review was appropriate for the problem. The researchers expressed interest in the overall problem and purpose, rather than focusing on a narrow area (usefulness, usability or performance) of research. This study is seen as important due to exploring an OA system's relationship to users, instead of a commercial or economic system's relationship to users.

## **Hypothesis**

Reformatting the reference question, the researchers do not form a distinct hypothesis. The research problem they seek to explore is which DL interaction attributes most significantly affect E-LIS usefulness and useability. This statement expresses a clear relationship between the DL evaluation and its users' opinions and needs. Although this statement is mostly based in theory, support is shown through previous studies in the individual areas of OA, information behavior and computer-user interactions, and DL evaluation methods. Initially investigated in the literature review, the variables were previously shown as test-worthy in relationship to other commercial e-print archive systems.

## **Method**

Developed by the Papatheodorou, Tsakonas, and their colleagues, the Interaction Triptych Framework (ITF) is a theoretical model that seeks to combine the knowledge and experience from information behavior and human-computer interaction. This method uses the DL components of system, content and user, which directly relate to the independent variables. The ITF was chosen to help define the relationships between users and system performance attributes on the performance axis.

All variables were clearly defined. The dependent variable chosen is the E-LIS OA e-print archive. The independent variables of usefulness, useability, and performance brought the study to a manageable level while still exploring this new user system. The three variables were initially studied independently,

and then compared in relationship to one another. Each one of these independent variables consisted of sub-attributes that were also explored independently as well as in relationship to the others.

The use of two distinct tools was employed to conduct the research. First, an online survey provided a way to define the E-LIS usefulness and useability. These thirty-four questions were split into three sections, with the first section addressing their familiarity and usage of E-LIS. Part two consisted of sixteen questions that covered the user's interaction with the database on usefulness, usability, and performance. These questions directly related to the concepts of usefulness, useability, and performance. The final twelve questions investigated the E-LIS functionalities in relationship to usefulness and useability. All questions were presented on a five point Likert scale. This standard question format is a common type in questionnaires. The researchers collected 131 individual valid questionnaires. Second, regression analysis was applied to the data in order to predict the attribute's strength and the impact of the system's functionalities. This analysis method is appropriate for this study since there is no previous knowledge about the processes that generate the data in an OA DLs. It also fits the data because it designed to test relationships between one dependent variable and one or more independent variables. The authors presented this methodology clearly and concisely in section 4.3.

Scoring the questionnaires used ratio levels of measurement. The reliability of the testing methods utilized Cronbach alpha tests (an internal consistency measurement) for each of the groups and individual items. The method scored above the 0.70 Nunnally and Bernstein threshold, with the scores presented in Table 1, thus giving reliability to the study.

### **Sample**

It is impossible to tell if the sample is representative of the population since the researchers used convenience sampling. To recruit participants for this study, the researchers posted notices on national and international mailing lists with the help of the national E-LIS editors. The study participants are drawn from volunteers of the E-LIS e-print archives, including the editors and administrative personnel. Although demographic data was collected from the respondents, the report displays no data to show that the collected data corresponds to the user population. This sample is also not representative of the population since participants were not randomly selected from all users, which could closely approximated by a database of users' IP address. Papatheodorou and Tsakonas make no reference to user selection or samples in similar studies; however, this is a common non-probable sampling strategy.

## Results and Discussion

Since another study of this nature does not exist it is difficult to relate the results to the literature review. The first section determined the characteristics of the participants. This section utilized quantitative methods to describe the group's characteristics. Almost half of the participants were unregistered users, over half resided in Europe, and the highest rate of introduction was via personal communication. The usage data for E-LIS possessed a normal distribution. In relationship to the user's participation level, the research found that the usage consistently increased as the role increased (*i.e.* from registered user to editor). Using a F-test for variance, considerable differences between the involved roles with their intense usage compared to the slight use of non-users. It is interesting to see the F-test (ANOVA) used to check variance since this explores the idea of one-sided probability of the likelihood that two variances are different. Other research is not available to which this study can be compared. The user data found in this section correspond to the literature review for similar studies in subject-based or institutional DLs on page 1236.

The second part of the questionnaire data showed overall positive attitudes to the thirteen attributes in the three categories of usefulness, useability, and performance. The only factor that showed incongruous data between roles is coverage. The mean responses differed considerably between editors and unregistered users. Using ANOVA, the high statistical significance of coverage did not correspond to all other sub-samples, which showed a probability of 0.001 ( $p > .001$ ). For the usability and performance questions, researchers followed the same format. In these two sections, the mean value and F-test was once again used to show statistical difference in each of the categories. In usability, the perceived ease of use and learnability increased as users' activity increased in the DL, reflected in the first three paragraphs of the literature review. The performance questions presented very high significant differences between frequent and infrequent users of E-LIS. The difference between this study and the information presented in the literature review is that this study focused on the human-computer process rather than the system-centered process. The data covered the same sub-areas of precision and recall. From these findings, Papatheodorou and Tsakonas placed the data into an inter-attribute Pearson Correlation table to show the correlations between user's opinions and the categories. Correlations are generally high (0.7) between each of the subcategories. From this table, the researchers can easily identify the attribute interactions that users most value. All the data presented to this point shows a logical progression towards answering the research question. To further assess the strength of each category, researchers applied stepwise multiple regression analysis to each category. This linear method first was checked against tolerance and this was found within acceptable limits, thus avoiding different data lines within the data. Next, a t-test to compare the independent data groups was applied. Both Relevance and Level were measured at the probability level of  $p < .001$ , while coverage was measured at the  $p < .01$  level. These three attributes

comprise over 65% of the variance. This showed the interaction between presented E-LIS usefulness and useability attributes.

The final part of the questionnaire presented data on the functionality and characteristics effect on the E-LIS OA system's usefulness and useability. The same methods were applied (quantitative questionnaire data and regression analysis). Once again, the analysis showed that the strength of the OA system lay in the perceived usability. This also supported the research questions assumptions of the interrelation between usefulness, usability, and performance. All of these areas were shown to affect user satisfaction.

In the discussion section, Papatheodorou and Tsakonas further relate their findings to the literature review. Their findings validated another study that also DL users' relationship and experience within the system affect their perceptions of the usefulness and useability. They also cited the need for further exploration into the roles and the concept of experience in the system since there were significant differences between editors and unregistered seldom users of the system. The conducted research also supported previous research in the relationships of ease of use and learnability, usefulness and the range of uploaded content, and system usage affect on users' preferences. This study also tested the theoretical foundation of ITF, finding that precision and recall are the main factors in forming users' opinions of performance. A paragraph devoted to the OA trend in the scientific community was provided, providing support for two related studies regarding structure retrieval tools as a high importance for users.

Papatheodorou and Tsakonas conclude their research statement with a call for e-prints and DLs to align themselves with user's work tasks and self-sustainable systems. They also see a need for further in-depth research in OA usefulness and useability to further support their conclusions in this area. Overall, the research results, discussion, and conclusion support their exploration into which DL interaction attributes most significantly affect E-LIS usefulness and useability.

## References

The reference list showed four articles published before 2000, with most of the remaining list displaying publication dates within the last five years. Most of the references presented came from scholarly journals. The content is comprised of some well-known names in the library and information studies field, including information presented at IEEE conferences and written by the Center for Research Libraries. Presented in ALA format, the list is complete and easy to navigate if the reader desires further reading.

### General Comments About the Report

Overall, the information presented is complete. However, the unbiased information feels to be condensed, and lead to some language clarity problems. To fully understand the literature and research presented, this reader had to read the article at least five times to grasp the general direction. The literature review seems to skip from the usability and usefulness of commercial or private systems, into digital evaluation methods, and then back into the usability and usefulness of OA systems. From the literature review, it is difficult to tell exactly where the literature review pertains to the research in the present study. This section should be more streamlined or better organized to make it more meaningful to the reader and show a better linear relationship to the study's methods.

The strength of this study came in the fact that it studied an OA system, a mechanism that is gaining popularity as an information delivery method. In my perception, the largest weakness showed in the vast exploration of different areas in OA user interaction and system attributes. Since it is one of the first exploring OA systems, it gave an umbrella view of this information delivery method that can now be built upon by other researchers. As a matter of convenience, the sample consisted of volunteers instead of a random sample. Also, instead of a questionnaire, interviews could have been done to acquire the participants' perception of E-LIS. This may have provided other insights into the users' relationship with OA repositories.

Overall, I feel that this study presented excellent data but had problems relating it to previous research until the discussion section. It is a new area that requires further in depth research to fully understand its functions, relationships, and implications on open access digital libraries and their users.