Lessons Learned from Observing User Behavior through Repeated Usability Evaluations

Hanmin Jung, Mikyoung Lee, and Won-kyung Sung

Abstract—Academic research information service is a must for surveying previous studies in research and development process. OntoFrame is an academic research information service under Semantic Web framework different from simple keyword-based services such as CiteSeer and Google Scholar. The first purpose of this study is for revealing user behavior in their surveys, the objects of using academic research information services, and their needs. The second is for applying lessons learned from the results to OntoFrame.

Keywords—User Behavior, Usability Evaluation, OntoFrame, CiteSeer, Google Scholar, Academic Research Information Service.

I. INTRODUCTION

ACADEMIC research information service is a Web site serving literatures such as papers, patents, and technical reports with search and browsing functions. It is a must for surveying previous studies in research and development (R&D) process. Several studies on usability evaluation with online academic libraries and academic Web sites were performed [1] [2] [3] [6]. However, they evaluated the services with simple keyword-based search tasks which are only a part of survey activities. For example, McGillis evaluated MUN (The Memorial University of Newfoundland) library Web site with restricted search tasks such as finding books, journal articles, and online helps [3]. Adlington, referring mainly the results of Jascó, failed to notice the ways considering R&D effectiveness because it deals with only basic components including menu name, duplication in search results, multiple links, and citation index on Google Scholar [1]. Nygren, as another study on Google Scholar, also focused on simple keyword search tasks for finding relevant papers about the user's thesis [6].

The first purpose of this study is for revealing user behavior in their surveys, the objects of using academic research information services, and their needs. The second is for applying lessons learned from the results to OntoFrame which is an academic research information service based on Semantic Web technologies.

II. USABILITY EVALUATION

We performed two usability evaluations. The first mainly

Authors are with Korea Institute of Science and Technology Information, 52-11 Eueon-dong, Yuseong-gu, Daejeon, 305-806, Korea (corresponding author to provide phone: +82-42-828-5087; fax: +82-42-828-5198; e-mail: jhm@kisti.re.kr).

consists of FGI (Focus Group Interview) to catch the needs of users about academic research information services. We found that the users want to acquire easily indirect information such as topic trends, relevant topics, topic–centric experts, and social network. The second takes aims at finding whether the needs can be satisfied on existing academic research information services or not. The evaluation with thinking-aloud protocol consists of two usability tests; the first is for observing user behavior in achieving direct information that can be acquired easily through simple keyword search, and the second is for indirect information that requires information analysis and inference. We performed them in comparative ways with Google Scholar (http://scholar.google.com/) and CiteSeer (http://citeseer.ist.psu.edu/).

We selected nine users (graduate students at KAIST [1]) participated voluntarily. It is expected to find over 95% of usability problems with the number of the users according to Nielsen's experiment results [4] [5]. All of them are using at least one of Google Scholar and CiteSeer for their surveys. Some of them are also using IEEE Xplore and Springer Web sites.

A. User Behavior in Acquiring Direct Information

The first goal requiring simple keyword search consists of three tasks as follows.

TABLE I
GOAL 1 WITH THREE TASKS FOR ACQUIRING DIRECT INFORMATION

1. Find papers of which titles include "neural network."
2. Find relevant papers for 'neural network' but their titles should not include "neural network," "neural," and "network."
3. Find papers about 'face recognition' based on 'neural network' methodology.

The users found papers easily using some facilities such as 'restrict to', sorting, and advanced search for the first task. However, the second task made different user behavior. The users read abstracts and full texts of the papers of which titles do not include 'neural network' with CiteSeer whereas they had difficulty in finding papers with Google Scholar because it does not provide abstracts and also shows paper titles including "neural network", "neural", or "network" as its search results within several pages. All of them queried with "neural network

World Academy of Science, Engineering and Technology
International Journal of Industrial and Manufacturing Engineering
Vol:1, No:7, 2007

face recognition" to find papers for the last task. However, some users failed when using CiteSeer because it provides poor search results in low precision. Google Scholar also has a problem related with inconsistency in that various orders of query keywords (e.g. "neural network" and "network neural") yield different search results.

Intuitive search results in high precision on Google Scholar would facilitate achieving the first and the third tasks whereas no provision of abstracts and metadata disturbs seeking for 'papers by topic' (the second task). Search results in low precision on CiteSeer makes considerable difficulty in finding papers efficiently. They wandered from page to page for gathering appropriate papers.

B. User Behavior in Acquiring Indirect Information

The second goal of increasing difficulty requiring information analysis and inference consists of four tasks as follows.

TABLE II
GOAL 2 WITH FOUR TASKS FOR ACQUIRING INDIRECT INFORMATION

1. Find relevant keywords and topics for 'neural network' (with grounds).
2. Find experts studying 'neural network' (with grounds).
3. Find relevant keywords and topics for the first ranked expert studying 'neural network' (with grounds).
4. Find researchers familiar with the first ranked expert studying 'neural network' (with grounds).

The above four tasks are for measuring the ability to serve indirect information like topic-centric experts and social network. High deviation between the users exists when compared with goal 1. For the first task, the users use their own background knowledge to extract some meaningful keywords from abstracts and titles in search results on both CiteSeer and Google Scholar. Some tried to search introductory papers to find relevant keywords for 'neural network'. However, the user behavior is quite different between the services in achieving the other tasks. Google Scholar provides 'All Results', a kind of researcher list (Fig. 1) whereas CiteSeer does not serve any researcher information.

The users utilize citation information for finding authoritative researchers when using CiteSeer for the second task. They click the first paper in search results, and then select the first author of the paper with the highest citation index. On the other hand, the users select one of 'All Results' with an involuntary action when using Google Scholar. They would be credulous about the information. For the third task, the users tried to extract some keywords from the abstracts and titles in the papers of the chosen author previously. But, many of them failed on CiteSeer because the quality of search results from person name queries is very low. CiteSeer would not permit search in author field. The users think that co-authorship is closely related with the familiarity between two persons, thus they chose co-authors of the export acquired from the third tasks as the answers of the last task.
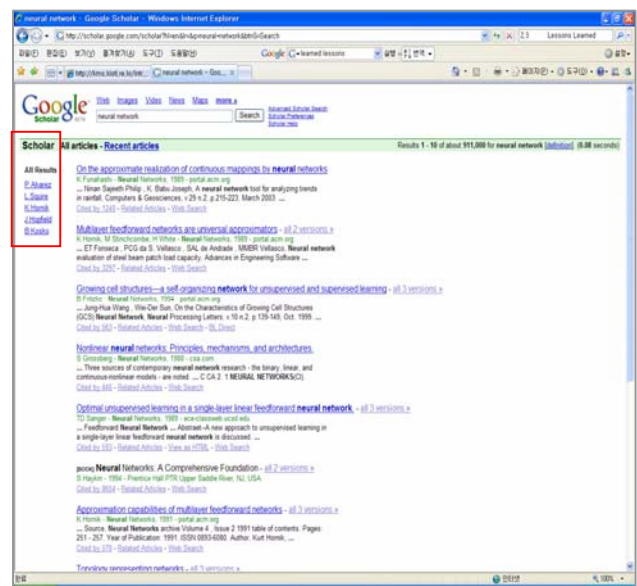

Fig. 1 'All Results' in search results on Google Scholar

In conclusion, the users often failed to perform the tasks with CiteSeer because it neither supports such service functions through standard approaches nor a search engine in satisfactory precision, which enforces the use of unreliable users' heuristics. Even in the case of success, high variation exists between the users. 'All Results', provided by Google Scholar, plays a decisive role in leading users to give their answers. These results indicate that an academic research information service had better prepare distinctive features for supporting various survey activities directly. Of course, they should be always reliable and understandable.

C. Interview Results

Opinions in italics back up the above conclusions. The users have suffered from acquiring such indirect information on CiteSeer and Google Scholar because the services are kinds of academic research information services under keyword-based scheme. The users are familiar with Google interfaces and have faith in 'All Results' with no doubt. In conclusion, they feel the need of advanced functions for facilitating surveys even though CiteSeer provides abstracts and citation information, and Google Scholar has a search engine in high precision.

World Academy of Science, Engineering and Technology
International Journal of Industrial and Manufacturing Engineering
Vol:1, No:7, 2007

TABLE III
PROS AND CONS OF CITESEER AND GOOGLE SCHOLAR

CiteSeer
Pros: citation information, BibTex (for latex users), multiple file formats
Cons: unfamiliar user interfaces, unclassified information, *difficulty in finding topics*, no highlighting in search results, low search speed
Google Scholar
Pros: familiar user interfaces, *'All Results'* (relevant researchers), high search speed, citation information, advanced search, reliable search results, search in date field
Cons: absence of sorting by citation, *insufficient functions for acquiring indirect information*, absence of abstract

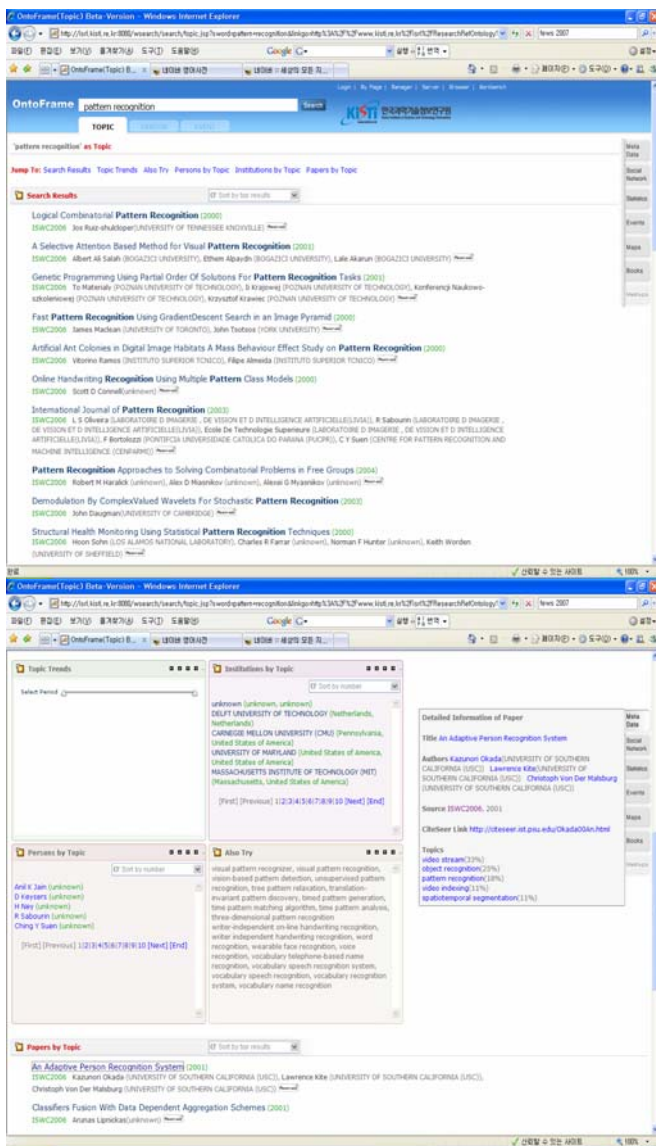## III. IMPLEMENTATION OF ONTOFRAME

Fig. 2 Example of topic page in OntoFrame service (upper and lower parts)

A. Lessons Learned

The precision of search engine should be reliable in order to prevent confused user behavior in direct and indirect information finding. The engine needs to support search in each field such as title search and author search. Consistent search results independent of the orders of query keywords also are obligatory.

The users ask for academic research information service to provide advanced functions for acquiring indirect information directly. Papers by topic, persons by topic, social network, and topics by expert belong to the helpful information. The next chapter describes how we put them into OntoFrame. First of all, they should be reliable because the users tend to have no doubt on correctness as shown through 'All Results' of Google Scholar.

An intuitive style of Google interfaces gives comfort to the users. Thus, user interfaces for basic functions like search results need to refer the style or those of similar commercial portals.

After the first evaluation, we modified the service goals and functions of OntoFrame; First, service functions were reorganized by entity pages such as topic, person, and event. For example, topic page consists of search results, topic trends, persons by topic, institutions by topic, also try (compositionally relevant topics), papers by topic, social network, and so on. Second, intelligent search logic was introduced. When the users query, OntoFrame automatically determine which entity page would be generated under URI (Uniform Resource Identifier) scheme. For example, user's query "neural network" leads topic page because it has been registered into ontology manager as one of topic instances. The search engine ensures high precision and search in every field, as well. As figure 2 shows, search results follow Google style whereas the other advanced functions are in movable boxes.

Currently, OntoFrame includes 114,337 papers gathered from CiteSeer Open Access Metadata. We refined them with identity resolution technology. Thus, 161,853 persons, 160,568 topics, 17,093 institutions, and 730,360 locations with GPS values were added. After implementing it, we performed a usability test with the same goals and persons involved in the second evaluation. Most of them achieved the tasks successfully, especially those of goal 2 in efficiency. They gave the same answers because OntoFrame provides such service functions directly.

IV. CONCLUSION

We performed two usability evaluations with focus group interview and usability test. From observing user behavior during the evaluations, we found that the users ask for academic research information service to provide advanced functions for acquiring indirect information directly. Search engine in high precision and user-familiar interfaces are obligatory for supporting the functions effectively. Finally, we introduced these lessons learned into OntoFrame. Our future works includes the ways to leverage use of our service.

REFERENCES

[1]   J. Adlington, Checking Under the Hood: Evaluating Google Scholar for Reference Use, In Journal of Internet Reference Services Quarterly 10(3/4), 2005.

[2]   B. Callicott, Google Scholar vs. Library Scholar: Testing the Performance of Schoogle, In Journal of Internet Reference Services Quarterly 10(3/4), 2005.

[3]   L. McGillis and E. Toms, Usability of the Academic Library Web Site: Implications for Design, In Journal of College & Research Libraries 62(4), 2001.

[4]   J. Nielsen, Card Sorting: How Many Users to Test, Jakob Nielsen's Alertbox, July 19, 2004.

[5]   J. Nielsen, Quantitative Studies: How Many Users to Test?, Jakob Nielsen's Alertbox, June 26, 2006.

[6]   E. Nygren, G. Haya, and W. Widmark, Students Experience of Metalib and Google Scholar, Dnr 63-612-2005 Report to BIBSAM, University of Stockholm, 2005.