

Personalized Information Retrieval Services for Digital Libraries

Hanumat G Sastry, Ph.D
Centre for Information
Technology
University of Petroleum and
Energy Studies, India

Venkatadri. M
Centre for Information
Technology
University of Petroleum and
Energy Studies, India

Lokanatha C. Reddy, Ph.D
Dept. of Computer Science
Dravidian University
India.

ABSTRACT

Digital libraries are huge and complex information systems. Digital library users are from various backgrounds with diversified information requirements. Digital library users are fed up with the information overload problem due its unsophisticated search features. Digital libraries can be effectively implemented by addressing the information overload problem. Hence, this paper presents newly designed personalized information retrieval services architecture and its implementation details to enhance the digital library users' search experience.

Keywords

Digital Libraries, search features, information overload.

1. INTRODUCTION

Digital libraries are huge and complex information systems, the contemporary developments in information and communication technologies have given a boom to the development of the digital libraries [1]. Digital libraries disseminate the information to the users across the world by crossing the physical barriers, various digital library projects have conceived in large scale in order to share the information/knowledge to the users in an easy and effective manner. Research has shown that digital library users' are from various backgrounds and their information requirements are diversified [2]. Contents in digital libraries are ever growing and most of the digital library systems are providing unsophisticated and monolithic search techniques without considering the user's background. Because, of these reasons digital library users' are fed up with the information overload problem [3][4]. To address the information overload problem in order to enhance the digital library search experience, this paper presents the newly designed personalized information retrieval services for digital libraries. This paper is organized as follows section 2 provides the literature review, section 3 provides the architecture and implementation details of personalized information retrieval services for digital libraries and finally section 4 presents the conclusion.

2. LITERATURE REVIEW

The information needs of individuals have long been studied by researchers in marketing, education, and information science. Various researchers have studied the information

needs and interaction behavior of the various types of users in information systems and digital libraries [5] [6] [7] [8] [9] [10]. Researchers have identified various levels of information requirements like specified visceral, conscious, formalized, and compromised levels [15]. In early days Z39.50 protocol was widely used for searching the libraries [11]. Search engine technologies have significantly improved in the last decade, Europeana, Apache's Lucene and Solar are the major search engine technologies for digital libraries [12] [13] [14]. In recent days, Users' profiling, visualization, personalized search tools have widely taken place in digital libraries [16][17], various researchers have worked to enhance the search experience as well as to minimize the information overload problem [3] [4]. However, research has shown that there is a large scope for improved search mechanisms/techniques to address the information overload problem of digital libraries.

3. PERSONALIZED INFORMATION RETRIEVAL SERVICES

To address the information overload problem in order to enhance digital library search experience, we have designed the personalized information retrieval services system as an add-on module to our semantic web enabled user interface system for dspace digital library [18]. This service is primarily depends upon the user profile and usage log. Due to this, the personalized information retrieval service module can able to filter the user interested documents from the huge digital library search space.

This service consists of several sub-modules like Personalized Information Retrieval (PRI) module and Information Filtering Tools (IFT) module etc. Personalized Information Retrieval (PRI) module provides the mechanism to retrieve the relevant documents to the user. This module retrieves the user relative documents based on the Information Filtering Tools with the help of Web 2.0 tools. Information Filtering Tools (IFT) module filters the documents based on the Knowledge Builder (KB) module. Knowledge Builder module represents the user interests in Semantic networks, these semantic networks are built in a supervised way, by integrating the User Model, Domain Ontology and Usage log. Information Filtering Tools (IFT) module constantly updates itself, based on the usage logs and related user tags.

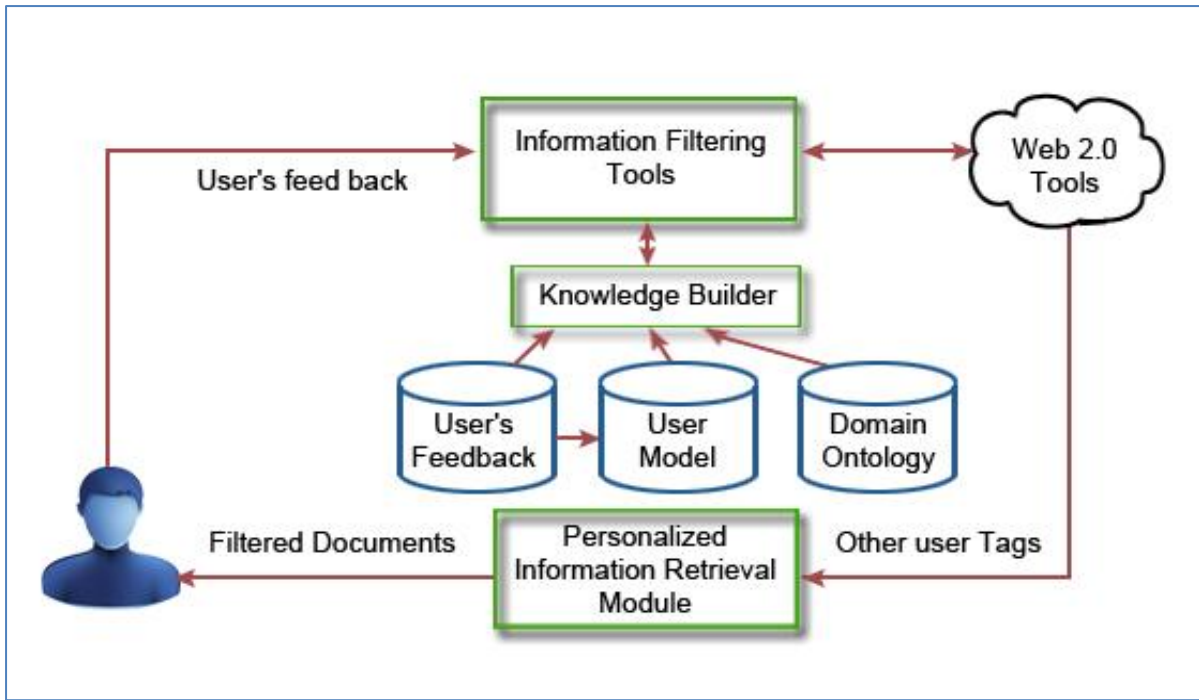


Figure 1: Architecture of Personalized Information Retrieval Services for Digital Libraries

The following screen shot shows the sample search result page of our system.

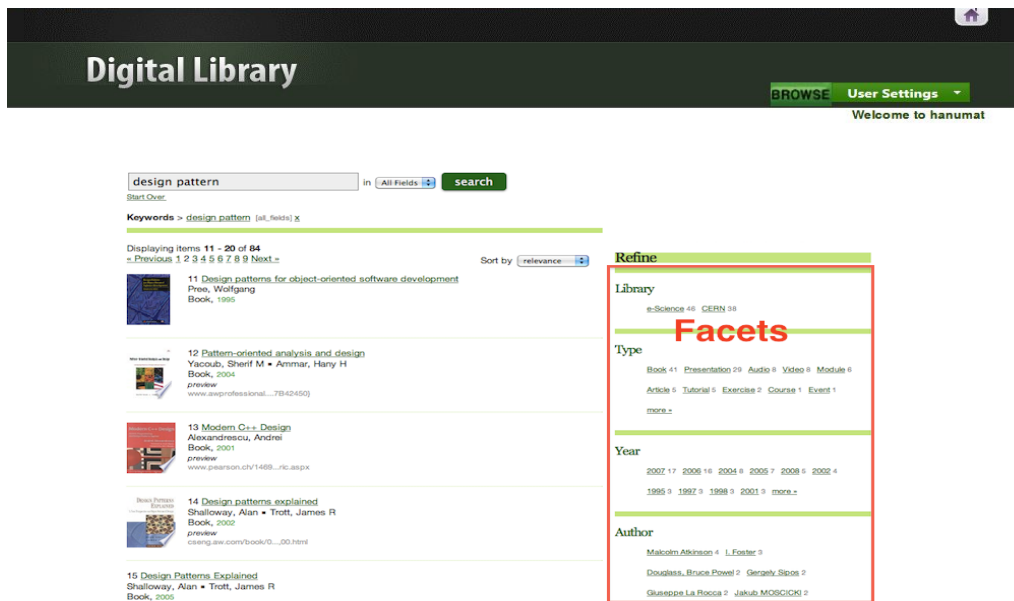


Figure 2: Personalized Search Results Page

4. CONCLUSION

In this paper we have presented our newly developed personalized information retrieval services for digital library to address the information overload problem. We successfully implemented this service in our prototype digital library system; the initial search results are encouraging. Further, we are planning to incorporate the data mining techniques on user

profiles and usage logs in order to improve the search mechanisms.

5. ACKNOWLEDGMENTS

The authors thank the University of Petroleum and Energy Studies, India and Dravidian University, India for the support extended to them for research in thrust areas of Computer Science.

6. REFERENCES

- [1]. Arms, W.Y. 2000. Digital Libraries. Cambridge. MIT Press.
- [2]. Hanumat G Sastry et.al. 2011. User Interface Design Challenges for Digital Libraries. International Journal of Computer Applications. Vol-15 No.6. Available at: <http://www.ijcaonline.org/archives/volume15/number6/1955-2616>
- [3]. A. Tejada-Lorente, C. Porcel, A. Martinez. G. Lopez-Herrera and E. Herrera-Viedma. 2011. Using memory to reduce the information overload in a university digital library. Intelligent Systems Design and Applications (ISDA'11).
- [4]. F. Blom. 2011. Information Overload and the Growing Infosphere: a Comparison of the Opinions and Experiences of Information Specialists and General Academics on the Topic of Information Overload. MSc. Dissertation, Uppsala: Uppsala Universitet. Available at <http://uu.divaportal.org/smash/record.jsf?pid=diva2:438746>
- [5]. Paisley, W. 1980. Information work. Progress in Communication Sciences. Vol 2(113-165).
- [6]. Wilson, T. 1981. On user studies and information needs. Journal of Documentation, 37(1), 3-15.
- [7]. Krikelas, J. 1983. Information-seeking behavior: Patterns and concepts. Drexel Library Quarterly, 19(11), 5-20.
- [8]. Marchant, M. 1991. What motivates adult users of public libraries?. Library and Information Science Research, 13(3), 201-235.
- [9]. He, D., Goker, A. 2000. Detecting session boundaries from web user logs. In: BCS-IRSG.
- [10]. Hshieh-ye, L. 1993. Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. JASIS 44.
- [11]. Lynch C.A. 1997. The Z39.50 information retrieval standard. D-Lib Mag. 3(4).
- [12]. <http://www.europeana.eu/portal/>
- [13]. <http://www.lucene.apache.org/core/>
- [14]. <http://www.lucene.apache.org/solr/>
- [15]. Gary Marchionini, Catherine Plaisant, and Anita Komlodi. 2003. The People in Digital Libraries: Multifaceted Approaches to Assessing Needs and Impact. Available at <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.64.2972>
- [16]. Stan Ruecker, Ali Shiri, Carlos Fiorentino, Amy Stafford, Mark Bieber, Matt Bouchard . 2011. Exploratory Search Interfaces for the Unesco Multilingual Digital Library: Combining visualization and Semantics. Journal of the Chicago Colloquium on Digital Humanities and Computer Science.
- [17]. Seonho Kim. 2008. Visualizing Users, User Communities, and Usage Trends in Complex Information Systems Using Implicit Rating Data, Ph.D thesis. Virginia Polytechnic Institute and State University.
- [18]. Hanumat G.Sastry et.al, 2010. Significance of Web 2.0 in Digital Libraries. International Journal of Computer Science and Engineering, Vol.2 No.6. Available at: <http://www.enggjournals.com/ijcse/doc/IJCSE10-02-06-97.pdf>