Web 2.0 features based personalization for academic digital libraries: an application for collaborative environment

Aditya M. Bhide[a], Yoo Jae Heung[b], and Rho Jae Jeung[c]

[a,b,c] School of IT Business, Information and Communications University
Daejeon - 305-732, Republic of South Korea
Tel: +82-042-866-6335, Fax: +82-042-866-6331, E-mail: {aditya[a], behappy[b], jjrho[c]}@icu.ac.kr

Abstract

Huge amount of useful information is getting added to the academic digital world everyday. This digitalized information is heterogeneous and remained limited to a local domain only. How to integrate and share this scattered knowledge with simplicity remained a big challenge. In this paper, we propose a prototype of web application for accomplishing above task using web 2.0 features; it makes traditional academic digital libraries to form as a special case of collaborative environment. This application enables users to customize their own resources to form their own virtual digital library. Further users can utilize personalized information of other users working in similar research domain to form true collaborative environment. The model and framework for the new look of academic digital library is presented.

Keywords:

Academic digital library, Collaborative environment, Customization, Personalization, Web

Introduction

Advancement in web technology has changed the way business is done over the Internet drastically. The e-commerce giant players like eBay.com, amazon.com used web personalization techniques effectively in their business activity for increasing their market share. Google, Yahoo [6] also started giving personalized home pages to their users. They treated web personalization as a strategy rather than technology to satisfy the needs of their own customers. Unfortunately, web is limited as a content publishing or browsing platform for most of the academic digital libraries. Academic digital libraries did not adopt the advancement in web technology, especially web 2.0 as much as it could have. There is very less attention paid towards using web to make academic digital libraries as collaborative environment from researchers' information needs point of view. In academia, there is a strong need for researchers to share their resources and thoughts for achieving rich quality of research work. In collaborative work, knowledge sharing is a key factor. Collaborative work can produce efficient outcome as one meets other users who are doing the same task but in better fashion than that of one is doing. In this regard, academic digital libraries can be treated as a classical example of collaborative information system if local knowledge can be integrated and shared effectively. This paper reduces this gap between researchers' information need and availability of knowledge using bridge of web 2.0 features.

In the next section we describe the personalization in library management systems and our approach towards it. The model and its flow are elaborated in section 3. Section 4 throws light on the framework of the web application for the deployment.
In final section, we conclude with outlining the future work.

Personalization in academic digital libraries and related work

Web technology can basically enhance any type of e-services on three major dimensions. They are user interactivity, content creation, and knowledge sharing. User interactivity addresses mainly user convenience. Studies based on Technology Acceptance Model (TAM) showed that perceived ease of use is one of the major determinants among all considered factors for academic digital libraries [12]. The content for academic digital libraries is high quality research database. This dimension makes it different from other business oriented information systems. But this rich content is only known to researchers in local context. Its value can be increased if and only if publication of this knowledge is shared and reused by other researchers for their work. It is cumulative process which in turn gives birth to further innovative work by others. Personalization can play active role in knowledge sharing dimension. Many times customization and personalization are treated as interchangeable terms. Customization is under direct user control where as personalization is driven by computers [9]. Customization is the way one arranges resources as per ones personal needs; where as personalization is finding out user patterns and making adaptive changes to our system. Ideal academic digital library should satisfy requirements of all these three dimensions.

Researchers are closely associated with academic digital library, a computer tool which opens door of knowledge for them. There are many academic digital libraries developed in recent time. Although they are under category of

personalized academic digital libraries; there main focus was on customization only. Habib with his conceptual library 2.0 model tried to make theoretical foundation of Library 2.0 in academic digital libraries [5]. Among other recent ones, University of Queensland built information system for their library; but its major focus is again on customization [3]. There are several Open Source Software (OSS) movements can be found especially for academic digital libraries as well [10]. Many university libraries [2, 4, 8, 13, 14] are developed based on versions of MyLibrary project. Firstly, almost all of them are built using static web 1.0. The other problem is that most of them focused on improving MyLibrary technically only. There more stress is on improving faster access, adding security, introducing book or course management, making e-mail or chat services inside academic digital libraries, making on-line help better, etc. Some of these enhancements are only from librarians' point of view or personal usage point of view. In all these enhancements only personal context of user in local environment is taken into consideration.

There is no way to find how one is extracting others knowledge and using it. Availability of such information is more vital than just browsing through collection of others work via academic digital libraries. Fewer efforts are put in this sense. One certainly does not expect improving digital library just as a local information system.
We tried to give digital libraries an extension where one can not just share his information but will also be benefited from others as well. Web 2.0 has potential to get academic digital libraries close to human behavior. Web 2.0 revolution introduced many features for web applications like RSS feed, wikis, IMs, better search, drag and drop facility and so on [7]. Features are the interface for wide variety of resources. Dynamic response achieved by web 2.0 can reduce the click distance in academic digital libraries. Our current work in progress takes one step further by incorporating web 2.0 features in academic digital library of our university.

There are several techniques available for web personalization [1]. There is a base criterion for every web personalization technique. Rather restricting ourselves only to customization, we took the approach ahead for combination of customization and personalization. The relative merits, opportunities, performance evaluation of this combinational aspect of customization and personalization are well described by John Russel [11]. The integration of customization and personalization is the most effective channel to enhance academic digital libraries as an information system.

The point to emphasis here is that it is more important of ``how'' rather than ``what'' researchers' access in academic digital libraries. With this, researchers save not only effort but also search time in acquiring and sharing their knowledge in easiest way with increased efficiency to form a perfect collaborative environment. In this way, the same academic digital library can be turned into an iterative and interactive feedback system where personalization is used for customization and customization itself is used to establish personalization with a perfect balance. This information system is our experience of web based learning technology where information can be retrieved, explored, manipulated and shared among similar peers. They indeed not just discover knowledge; but enhance it with their own research for group of other similar researchers in united way.

Basic model

The basic model for our system is user centric model.
Data is represented in form of features is a concept of this model. Let we have a universe set of n users as $U = \{u_k : 1 \leq k \leq n\}$. Every user is associated with two profiles namely $<p_a, p_b>$ such that 'p_i' is x-dimensional vector $(a^i_1, a^i_2, \ldots\ldots, a^i_x)$. '$p_a$' is termed as primary profile which is actual user profile for user u_k that reflects user's real interest in research. 'p_b' is an alternative or secondary profile which closely matches user's real interest in research. These two profiles give user a flexibility to switch to any profile at give point of time. With these two profiles user can change over to any one of the profile easily to get customized information of other users having similar profiles. Let there be universe set of m web 2.0 based services $S = \{f_j : 1 \leq j \leq m\}$ offered by the system. 'f_j' is a y-dimensional vector $(c^j_1, c^j_2, \ldots\ldots, c^j_y)$. We called this as component vector. Thus, every service is associated with its own components.
We have matching function '$\mu_{pa}{}^{uk1, uk2}$' which compares 'p_a' for any two users 'u_{k1}' and 'u_{k2}'. Let t be the threshold value for above matching function. If the value delivered by comparison is above threshold level t then it signifies that these two users share similar interest in research else they do not. This test will identify what we call as test of potential similar users (PSU).

$$\text{Is } u_{ki} \text{ PSU? } \begin{cases} 1 \text{ If } '\mu_{pa}{}^{uk1, uk2}, \geq t \\ \text{else } '0' \end{cases} \quad (1)$$

We can define function μ_{uk} as

$$\mu_{uk} : S \rightarrow [0, 1] \cup \perp \quad (2)$$

where $f_j = \perp$ indicates that 'f_j' is not getting used by that particular user, 'u_k'. Also $f_j \in S$; let '$S_k^{(nr)}$' be the set of features not recommended by user 'u_k'.

$$S_k^{(nr)} = \{f_j : f_j \in S \wedge \mu_{uk}(f_j) = \perp\} \quad (3)$$

In the same fashion, we can have '$S_k^{(r)}$' as a set of features recommended by the user 'u_k'.

$$S_k^{(r)} = S - S_k^{(nr)} \quad (4)$$

Hence, the final strong and weak set of recommended features from all users can be $S^{(Rs)} = \cap S_k^{(r)}$; $S^{(Rw)} = \cup S_k^{(r)}$.

Features and flow

The logical hierarchical diagram for the above mentioned model is shown in figure [1]. The direction from top to bottom represents customization where user performs actions; whereas direction from bottom to top represents personalization recommended by system.

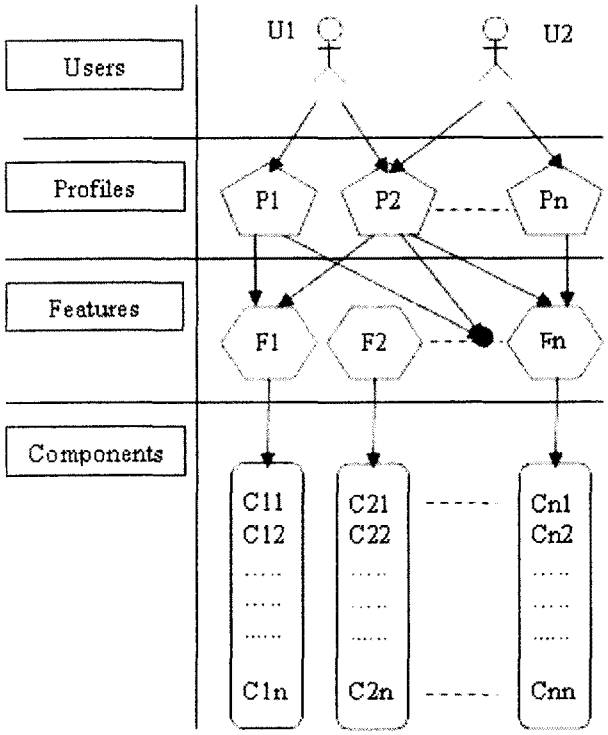

Figure 1 - Logical hierarchical figure for model

As a first step, every user is registered via authentication module. For interaction any user agent like browser can be used. These users create their own two profiles matching their areas of interest. As shown, user 'U_1'and 'U_2' can share same '*P2*' as their primary and secondary profiles. Based on these profiles, they are offered features by the system. There features are the same which are currently getting used by other users in the same knowledge area. Given the set of users, the estimation of these features is done as per the mentioned model. This assistance from the system will allow users to discover how other similar potential users are employing their resources. Further, users are now free to customize features as per their requirements. Their customized environment is saved as an association with their own profile. This new information is used by system for recommendations to new coming naive users.
With more active users eventually it will give supportive environment for all users. With the customization and personalization going hand in hand, it will make the collaborative environment richer and richer.

Imagine there is a researcher who defined his profile as interest in area 'R_q' in local environment. Let there be other PSU in outside local environment. This PSU gained the expertness in research area 'R_q' over the time. His or her vast experience will be very much useful for others who are looking forward for resources in this area. It is very much time consuming process to locate right resources for new user. This PSU has a feature where he gets RSS feed from rare journal in area 'R_q'. Now knowing early of this RSS feed will help user who is seeking this information to start his research with. System will compare and provide strong and weak set of recommended features with soft components to new user. Knowing several web 2.0 features is trivial task for many users; but obtaining components associated with it is crucial. User can fine tune these acquired knowledge components from origin. These modifications can be considered by system for later recommendations to others. Many other features with their components can be retrieved as well and shared to leverage converging collaborative work environment.

Conceptual framework

The framework has three tier architecture as shown in figure [2].

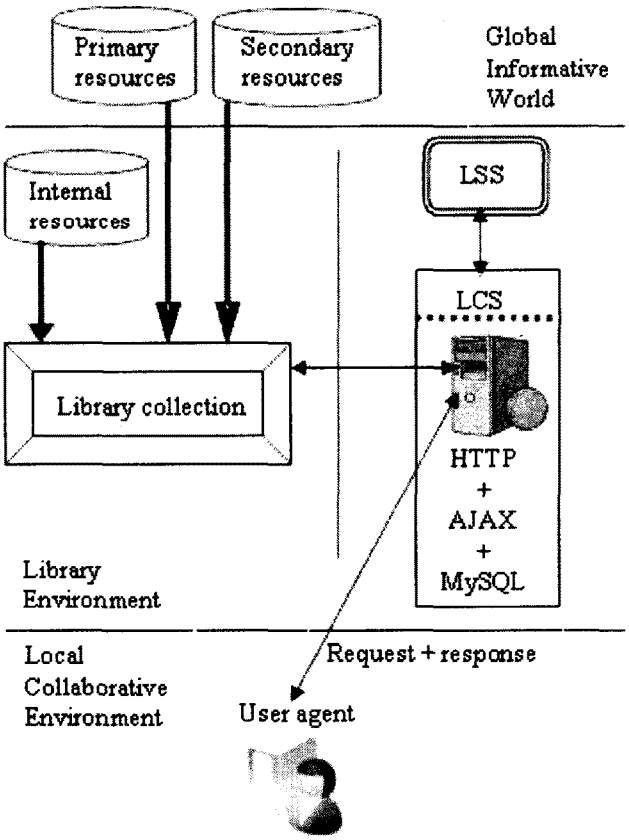

Figure 2 - Basic framework

The lower part is a local collaboration environment where all researchers are closely associated in same physical location. They can share their information and resources easily. The top part is a global informative world. Here

many useful information and resources are residing which are out of reach of local environment. Middle zone of library environment acts as a medium to make these two environments together to form complete set of collaborative environment. Library collection is not comprised only for internal resources but also fetches primary and secondary resources from external world. Internal resources are in-house features and information. Primary resource is a group of similar other outside local environments. Eventually, these all local environment systems will be integrated to form a system of systems. By secondary resources, we mean not the features directly but those components which are accessible by common features. Library servers are divided in two major parts. Library Core Server (LCS) serves typical services necessary to run research library. We built simple HTTP servers like Apache to serve users' HTTP requests, MySQL database server and so on. Library Support Servers (LSS) is a fundamental server of system. It acts as agent, who gathers, stores, manipulates and delivers the information as per functions described in the model. LSS is transparent to end user; and makes system to appear as an agent based recommender system. With technology like Ajax, the realization of this system was very much feasible for us. The added advantage of this application is that it is platform independent as being web based.

Challenges

The proposed system is influenced by Open Source activity. Our current library management system is a proprietary product. The convergence of these two efforts can hinder the progress. Secondly, there are two types of categories of users. One which experiments too much with personalization; and others take system as it is taken. For effective usage of feature based personalization, we are working on appropriate demonstration and manual preparation. Privacy issues are always a major concern in personalizing applications. Our other engineering group is focusing on privacy issues currently; and making it worth to mention in this paper as an open challenge. Also, we experienced cold start or latency problem with this web based application at least in initial stages as expected. We are trying to pay Special attention towards it as well.

Conclusion and future work

Quality of academic digital library usage can be enhanced using web personalization. In this paper, we delivered the features based basic model and framework for its deployment. It can be regarded as the library of 21st century with convergence of customization and personalization forming effective collaborative environment. Features are at the core of this concept of personalization. The developments of new features can be done independently and incorporated in the application easily at any point of time.

After gaining confidence with implementation of system with very basic features of web 2.0, the application should not be constrained only for internal university usage. We are planning to survey very soon to determine more features researchers expecting in the academic digital libraries. It will give us a critical clue for making this system robust and for more wide range of audience.

Acknowledgments

This research was supported by the MIC (Ministry of Information and Communication), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Advancement) IITA-2006-C1090-0603-0014.

References

[1] Sarabjot Singh Anand, and Bamshad Mobasher. (2005). "Intelligent Techniques for Web Personalization," *Lecture Notes in Computer Science, Springer Berlin / Heidelberg*, Vol. 4155, pp. 247-264.

[2] Cohen, Suzanne, John Fereira, Angela Horne, Bob Kibbee, Holly Mistlebauer, and Adam Smith. (2000). "MyLibrary Personalized Electronic Services in the Cornell University Library," *D-Lib Magazine*, Vol. 6(4).

[3] R. Foxlee, and K. Borchardt. (2000). UQ Library's First CIS: Towards Customising Information Services at the University of Queensland Library," *Proceedings of the 21st annual International Association of Technological University Libraries (IATUL) conference*, Vol. 10.

[4] James Ghaphery. (2002). "My Library at Virginia Commonwealth University," *D-Lib Magazine*, Vol. 8(7).

[5] Michael C. Habib. (2006). "Toward Academic Library 2.0: Development and Application of a Library 2.0 Methodology," M.S. thesis, School of Information and Library Science, University of North Carolina",

[6] Udi Manber, Ash Patel, and John Robison. (2000). "Experience with personalization of Yahoo!," *Communications of the ACM*, Vol. 43(8), pg. 35-39.

[7] Jack M. Maness. (2006). "Library 2.0 theory: Web 2.0 and its implications for libraries," *Webology*, Vol. 3(2)

[8] Tim McGeary. (2005). " MyLibrary: the library's response to the campus portal," *Online Information Review*, Vol. 29(4), pp. 365-373.

[9] Jakob Nielsen. "Personalization is Over-Rated", Accessed at http://www.useit.com/alertbox/981004.html in October 1998.

[10] Steve Oberg. (2003). "Open Source Software: An Introduction from a Serialist's Perspective," *Serials Review, Elsevier Science*, Vol. 29(1), pp. 36-39.

[11] John Russel. (2003). "Making it Personal: Information that Adapts to the Reader," *Proceedings of the 21st annual international conference on Documentation, ACM Special Interest Group for Design of Communication*, pp. 160-165.

[12] James Y. L. Thong, Weiyin Hong, and Kar-yan Tam. (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*, Vol. 57(3), pp. 215-242.
[13] Anamarija Rozic-Hristovskii, Iztok Humar, and Dimitar Hristovski. (2003). "Developing a multilingual, personalised medical library portal: use of MyLibrary in Slovenia," *Program: electronic library and information systems*, Vol. 37(3), pp. 146-157.
[14] Wu He, and Demei Shen. (2005). "MyPDL: a web-based personal digital library," Proceedings *of the 5th ACM/IEEE-CS joint conference on Digital libraries JCDL '05*, pp. 380-380.