Wondwossen Mulualem Beyene

Personalization and Usage Data in Academic Libraries: An Exploratory Study

Master thesis
International Master in Digital Library Learning
2012

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the program of International Masters in Digital Library (DILL) that gave me the wonderful opportunity of being part of the program as a student. It is the most wonderful opportunity I ever had in my life, that really made a huge difference not only to my knowledge as an information professional, but also added the most precious life experience of studying in three universities in Norway, Estonia, and Italy with students from United States, Cuba, Norway, Denmark, Italy, Poland, the Netherlands, Romania, Turkey, Azerbaijan, Bangladesh, Malaysia, China, the Philippines, Zimbabwe, Pakistan, Lithuania, Estonia, Iran, Spain and Uzbekistan.

My sincere gratitude and utmost respect goes to all my professors who taught me at the three universities and also the DILL administration that made it all possible. Thank you very much. I am so grateful for the privilege and words fail me to express my gratitude.

I say thank you to my supervisor Professor Nils Pharo for his most valuable advice and guidance in the course of this research. I would also like to extend my thanks to all who have cooperated in contribution of information that helped the realization of this thesis.

Finally ,I say thank you , Lord, for this wonderful opportunity.

ABSTRACT

Personalization is a service pattern for ensuring proactive information delivery tailored to an individual based on learned or perceived needs of the person. It is credited as a remedy for information explosion especially in the academic environment and its importance to libraries was described to the extent of justifying their existence. There have been numerous novel approaches or technical specifications forwarded for realization of personalization in libraries. However, literature shows that the implementation of the services in libraries is minimal which implies the need for a thorough analysis and discussion of issues underlying the practicality of this service in the library environment. This study was initiated by this need and it was done with the objective of finding answers for questions related to library usage data, user profiles and privacy which are among the factors determining the success of personalized services in academic libraries. With the aim of finding comprehensive answers, five distinct cases representing different approaches to academic library personalization were chosen for thorough analysis and themes extracted from them was substantiated by extensive literature review. Moreover, with the aim of getting more information, unstructured questions were presented to the libraries running the services. The overall finding shows that personalization can be realized in academic libraries but it has to address issues related to collecting and processing user/usage data, user interest management, safeguarding user privacy, library privacy laws and other important matters discovered in the course of the study.

Table of Contents

Tables

Figures

Chapter One

Introduction

1.1. Background

The world of global education is shaping up to be more student-centered, interactive and dynamic, enabling group work on real world problems, allowing students to determine their own learning routes and emphasizing competencies like information literacy to support lifelong learning (UNESCO, 2003). The growth of internet has catalyzed growth in e-learning and trends in education such as the Bologna process have put the learner at the center where resources and activities should be arranged according to each learner (Bologna Declaration, 1999, Ferran et al, 2007). This new scenario necessitates personalization of learning according to each learner's preferences, competency and so on (Ferran et al, 2007). Needless to say that libraries should mirror developments in their parent institutions, such developments coupled with the increasing popularization of academic libraries as learning commons give the libraries a chance to be centers of learning and content for their parent institutions (Baily, 2008).

The challenges academic libraries face are manifold. First, as mentioned above, they are affected by any political, social, economic or technical changes affecting their parent institutions. Second, they have to deal with user push and technology pull , review the way they do business and align themselves to their environment's needs which too are affected by technological changes. For instance, the immense potentials of communication and interaction made possible by web 2.0 have affected several professions including librarianship with implications for user expectations and the future of academic libraries (Coelho, 2011). Libraries have to thrive to satisfy the needs of their increasingly tech savvy users who expect their library to be up to the standard in relation to other information systems. There are calls for libraries to be proactive in their service delivery but being proactive requires good knowledge of users and their preferences which by itself is a challenge, given the rigorous privacy requirements libraries have to work with.

As libraries strive to deal with technology and changing user needs, they confront yet another big challenge: the information glut, which is a big threat for delivery of the right information to the right person at the right time. Libraries have to match user expectations in terms of content discovery, ease of use and speed of response otherwise they will be relegated to secondary role in favor of other information services (Coelho, 2011). There are even studies showing that, in the academic context, searching online catalogs is being abandoned in favor of simpler functionalities provided by internet search engines (Calhoun, 2006; Coelho, 2011). There are also researches showing a decreasing trend in the use of academic library websites and increasing use of Internet search engines and blogs (Kaur 2009; Chua and Goh 2010).This should be a wakeup call for academic libraries to retool their services, add value to their traditional services and reclaim their role as primary information service providers in their institutions.

The emergence of digital libraries has helped libraries to deal with information overload and problems it carried along. As explained by Smeaton and Callan (2005), the first generation of digital libraries were developed with relatively homogenous and well-informed groups in mind and that helped tailoring content to the needs of the right communities . Digital libraries played the role of content-to-community mediators by working on four areas such as content pre-selection, content structuring, content enrichment, and library services such as retrieval, access and annotation services (Neuhold et al., 2003).

However, as heterogeneity increased in terms of materials available and types of users served, the need for broader models for better understanding of users and provision of tailored services became apparent (Smeaten and Callan, 2005). This has made the global “one size fits all” approach, for instance, providing one interface to all users, a major weakness of digital libraries (Neuhold et al., 2003). This entails the need for collecting information on individual differences such as task commitments, level of competency, preferences and individual conceptualization of the resources for narrowing the gap between contents and users (Neuhold & et al., 2003).

The need for personalization was affirmed by joint NSF-EU DELOS Working Group Report as “*Digital libraries that are not personalized for individuals will be seen as defaulting on their*

obligation to offer the best service possible" (Smeaten and Callan, 2005, p.299). In the academic library context, personalization helps to map the right resources to the right users. It enables users to carve out a learning space for themselves out of a bigger library system and, going back to the roots of librarianship, we can say personalization helps librarians in fulfilling at least three of the five Laws of librarianship, i.e. "every user his/her book", "every book its reader" and "Save the time of the user".

However, despite the promises and potentials mentioned above, personalization doesn't top the list of important functions in many academic libraries. As the discussions in the following chapters show, there seems a gap existing between proposals or recommendations and actual implementations on the ground. This research paper explores the situation by making extensive literature review and through analysis of selected cases. It aims at rediscovering the contribution of personalized services in improving users' experience and investigating the issues determining success or failure of such services.

1.2. Statement of the Problem

Much of the research and development activities on personalization were done to create new technologies, understand the issue of personalization from business perspectives and develop novel technologies (Adomavicius and Tuzhilin, 2005). There have been top-down (e.g. ontologies, subject headings, user modeling) and bottom-up (e.g. data mining, text mining, web mining) approaches used to propose better recommender and personalization systems for digital libraries (Ferran et al., 2005). However, discussing the issue in practical terms would illustrate how complicated it is. For instance, the idea of personalization seems to negate privacy in libraries where privacy of users is sacrosanct. For example, all recommender systems require user data to run effectively. They need a critical mass of data on users' preferences, usage patterns, likes, etc. The greater the amount of data, the more robust a recommender will be. This puts libraries in an interesting position if their policies don't allow them to store user and usage data. In addition to this, the availability of tools, resources, competence, attitude of librarians, or the actual need for such services would help us to get the picture as how practical the services can be to academic libraries.

Studies show that the amount of research done on the actual implementation of such systems in academic libraries is minimal (Neilson, 2008). For example, according to the literature review made by Park et al (2012) on 210 research papers written on recommender systems in 125 MIS[1] journals between 2000 and 2010, the majority of them were found to be related to movie, shopping, image and TV programs. Most of these research papers appeared on journals dealing with expert systems applications (Park et al, 2012).

A longitudinal study performed on Portuguese university libraries between 2008 and 2010 on their use of Web 2.0 technologies shows that only three portals were found with personalization functionalities (Coelho, 2011). Adding one more research on academic library websites, Liu (2008) examined practices of 111 ARL member academic library websites and found out that only four of them have provided personalized library spaces called "My Library," "My Personal Library," or "My Search Space," . In that article she recommended the future academic library websites to include personalization in their design and proposed a conceptual model for library websites. The model included personalized spaces where a user can save and access preferred library resources, interact with others using tools such as instant messaging, e-mail, and bulletin board; work with data/documentation online, create and share content (using blog, wiki, or podcast), and integrate school/work activities (Liu, 2008).

Wakeling et al (2012) made a study on OPACs[2] in the UK and found out that, from 118 library OPACs, only 13 of them feature recommendation. Even out of these, eight of them were results from Encore discovery tools which discover new items in the library related to the searched item. These tools follow purely content-based approach that relies on existing metadata (such as keywords and subject areas). Only two of the OPACs were found to offer recommendations based on some form of collaborative filtering technique (Wakeling et al, 2012). One of those two OPACs is found at the Huddersfield University Library and it is one of the cases chosen for analysis in this paper. In another study made on 260 library OPACs in the US and Canada, Yang and Hoffman (2010) found out that none of the OPACs exhibited the

[1] MIS: Management Information Systems
[2] OPAC: Online Public Access Catalog

Amazon model of “those who borrowed this also…” recommendations. However, they said that 30 to 36 percent of those OPACs showed the “Spirit of that feature”, recommending items based on item similarity rather than similarity between tastes of users. The recommendation expressions used included “show similar”, “nearby items of the shelf”. “Similar works” etc. (Yang and Hoffman, 2010).

McLaughlin (2011) made a study about the usage of personalization features on some journal databases university libraries subscribe to. She found out that their usage was low and recommended for them to be marketed or incorporated as functionalities that would work without being noticeable.

To sum up the above discussion, there seems a gap existing between researches and implementations. The actual implementation of personalized services in libraries seems to be minimal (Lynch, 2001). One might ask here, “why”? Where is the promise of personalization to libraries? Can it really be useful or applicable? This is what this thesis tries to find out.

1.3. Research Objectives

There have been efforts being made by few academic libraries in form of enhancing OPACs with recommender tools and /or providing personalized information spaces (such as MyLibrary, MyShelf, etc.) on library websites. The aim of this research is to explore and analyze such cases and, by supporting the findings with literature review, present a perspective on how research and practice in personalization can be approached in the context of academic libraries. It puts more focus on the exploitation of user/usage data and their implications for personalizing information services. To this end, this paper aims at answering the following questions:

- Which functions of academic library can be assisted by personalization /recommender systems and what is the best way for implementing the services with the maximum possible convenience for the user?

- What does the usage trend of personalization features in academic library systems look like?

- User profiling: who does the profiling? How can it be done in digital environment? How can it be comprehensive and at the same time be in par with privacy requirements?

- Usage data: how can it be collected and be used to ensure better personalized services?

- What is the role of academic libraries in the design of personalized services?

1.4. Scope of the Research

As personalization is a generic term for individualization of products and services and in many literatures the term is mentioned along recommender systems, this section attempts first to show the definitional scope used in this study.

Smeaton and Callan (2005) define personalization as the way in which information and services can be tailored to match the unique and specific needs of an individual or a community by adapting presentation, content and service to the user's history, background, device, information needs, location, etc. They regard recommender systems as a particular type of personalization that learn about user interests and proactively identify and recommend information that meets their needs.

Neuhold et al. (2003) add that recommender systems follow information 'push' approach whereas personalized information access fall under the category of information 'pull' process. However, they too described recommendation systems as a type of personalized information access. Perugini and Gonçalves (2002) view personalization as tailored information service delivered to users in three ways:

1. Recommendation services.
2. Personalization of information access; which involves techniques and methods for the restructuring of information architectures to meet personal preferences, needs, or requirements.
3. Induction, exploration, and exploitation of social networks that can bring together scholars with common research interests

The term personalization is understood in this thesis in the sense discussed above. It' s scope is limited to discussing the role of recommender tools and personalized information spaces in enhancing user experience at academic libraries, putting more focus on utilization of user and usage data. It focuses on five purposefully selected cases representing different approaches to academic library personalization and attempts to get themes that can later be enriched by the literature review in order to get comprehensive answers to the research questions.

It might perhaps be necessary to say little about digital libraries as most of the literature consulted on this thesis invoke the term 'digital library' while discussing library personalization. As Tedd and Large (2005) explained, scholars from the field of librarianship and computer science seem to have competing definitions for this term. Scholars from the computer science community give much attention to its database and retrieval aspects whereas those from the library community view digital libraries as extensions of their traditional tasks which is selecting, collecting, organizing, conserving, preserving and providing access to information yet in another (electronic) form. The latter expression seems to reflect the reality in today's academic libraries and discussions on digital library personalization are understood in that sense.

1.5. Limitations

Due to time constraint and shortage of related works, the author admits this paper may not be as comprehensive as it should be. The data collection on the selected cases has relied heavily on examination of primary information resources such as original reports and blog posts and also open ended questions sent through e-mail. In most cases, the reports and blog

posts reviewed were recommended by the respondents running the services and also were written by themselves. Therefore, I believe this gives some degree of authoritativeness for this research. The other problem was, as mentioned in the following chapter, the difference level of maturity the concept of personalization has among the libraries. Identifying the right person to answer the research questions and attempting to create communication was constrained by time limitation.

1.6. Organization of the Paper

This research document is organized under seven chapters. The first chapter provides introductory information about the research problem, and organization of the paper.

The second chapter presents the methodology used to answer the research questions including the research framework and method of data collection.

The third chapter presents review of literature that provides conceptual base for the research and touches upon issues such as recommender systems, personalization of academic library portals, user profiles, usage data, privacy, technical and non-technical ways to protect user privacy in personalized environments and other related issues.

The fourth chapter presents cases that could serve as best examples on how the issue of personalization can be approached in academic libraries. In this chapter, the case of Huddersfield University's homemade OPAC, BibTip recommender which was originally integrated to Karlsruhe University Library OPAC but now being used at more than forty institution's OPACs, PORE (Personal Ontology Recommender) which is developed at National Chung Hsing University of Taiwan and was being used there until recently, and Ex Libris' bX recommender system which is integrated with Ex Libris products and also available for subscribers. In addition to the above four distinct examples of recommender services, the case of MyLibrary as a form of personalized library portal is presented.

The Fifth chapter comprises the analysis part of the paper and analyzes information from the previous two chapters to answer the research questions and issues related to them. It extracts

themes out of the cases analyzed, substantiates them with the literature review, categorizes and presents them under the research questions they answer. Chapter six provides conclusion, summarizing the answers for the research questions followed by chapter seven, which elaborates on some of the themes explored in the course of the research and provides future research directions.

Chapter Two

METHODOLOGY

2.1. Research Framework

As explained above and as reflected in the research questions, the main purpose of this paper is more on gaining deeper insights, understanding and perspectives on how academic libraries can view or approach the issue of personalization. In general, this research follows the qualitative research paradigm because, as explained by Hewit-Tylor (2001), this type of research is suited to show the reality of the area under investigation and enhance understanding of the situation by gathering as much data as possible from variety of resources. Therefore it suits this research's objectives.

Under the qualitative research paradigm, the methodology chosen for this purpose is the case study method as it is best suited for holistic and in-depth investigation of research interests (Feagan et al, 1991). Case studies can be exploratory, explanatory, descriptive, instrumental, intrinsic, and collective, (Tellis, 1997). This research is exploratory in one way because this method is appropriate for gaining deeper insights and familiarity for such themes where there are very few studies to refer to [3] or to explore areas where there is little theory available or measurement is unclear (Kohn, 1997). As this research depends highly on documentary information resources such as primary literature and data, the exploratory method was found to support the framework[4]. It can also be considered as collective case study because it aims at analyzing five distinct cases to get as much answers as possible for the research questions.

The first reason that forced choice of the case study method was the preliminary investigation I made and the difference in maturity of the concept I observed among libraries. Before the launch of this research, I interviewed and corresponded via e-mail with couple of librarians heading university library digital services section to check their idea about personalization , management of user data and profiles, and also the utilization of features such as MyLibrary,

[3] http://libguides.usc.edu/content.php?pid=83009&sid=818072
[4] http://en.wikipedia.org/wiki/Exploratory_research

e-shelf, etc that are available on their library systems. What I found out was that the concept of personalization was not that popular or a thing that they currently are giving serious attention. The next attempt was to search the Internet for academic libraries that are running a kind of personalized service. I was able to contact two libraries which, according to their websites, were running the services. One of them replied that they are at the earliest stage and plan to improve the website in the future so they cannot say anything about it. The other one directed my inquires to a company that helped them in designing their system. In addition to this, I followed the lead from literature review to contact university libraries which have OPACs enhanced with recommender tools. Most of those libraries were users of the Ex-Libris library system and the answer they gave me was that the system is still new to their library and the recommender functionality is not yet promoted.

Therefore considering such cloud of uncertainty and the time available for the research, it was found better to select and focus on relatively more mature personalization attempts, explore them as much as possible and compare the results with conceptual literature in order to answer the research questions and, by doing so, explore issues and challenges surrounding implementation of personalized services in the academic library environment. The case study approach was found to be the better option.

The second reason for the choice of the research method discussed above can be attributed to the shortage of related works to refer to. Literature is not short of explanations concerning the potential benefits of tailored information services to libraries (especially digital libraries) and there are lots of technical specifications and proposals written mainly by the computing society. However, as mentioned in the previous sections and also discussed in the following chapters, there is a serious shortage of study done on how those technical proposals are being implemented in libraries. Such studies discovered at the time of this thesis were mostly usability studies, for example, testing a recommender tool by asking participants to sit down and test it as per the instructions given by researchers and later ask them what they feel about the system (copac, 2011; Thomas et al ,2011); checking trend of usage by observing and checking statistics of registered users at some intervals in a given period and use the results to get a picture for the tool's usability (Ciccone,2005; Shedlock et al 2010), etc. The other type of study observed was about examining academic library websites and check for

presence of personalized features (Liu, 2008) and checking library OPACs for the presence of recommender functionalities on them (Wakeling et al ,2012) .

Therefore, instead of following similar trend and possibly repeat similar findings, I believed it would be more meaningful to take this research to a bit higher level in a way that summarizes those findings and also provides broader perspective and a kind of conceptual foundation for other researches to come. Therefore making in-depth investigation of the relatively mature cases and supporting the findings with extensive literature review was found to be the better way for conducting this research.

To provide a conceptual base for this research, search was made on the Web of Science database for articles containing topics on “personalization” and “University Library” or «academic library” and “recommender systems” and “academic library” or «University Library” which yielded total of 22 hits out of which 14 were from library and information science, 13 from computer science and the rest from other fields (there seems that some articles are categorized under more than one subject area). To get more results it was necessary to modify the combination to “personalization” or “recommender systems” and “library”. This time a total of 99 hits out of which 70 were from computer science field, 40 from the field of library and information science and the rest from engineering and operation research management science fields were found. As this research is more interested in answering those conceptual questions rather than dealing with the technical intricacies surrounding the research theme, the highly technical papers were avoided and 59 of those articles were used in the literature review. Other articles found via Google Scholar have also been incorporated among those discovered through the Web of Science. All in All, 117 documents were consulted in the course of this research.

Then, following the lead from the literature review and hints from the e-mail correspondences, the OPAC of Huddersfield University Library, BibTip Recommender system which started its service at Karlsruhe University Library but currently being used by 47 more Libraries , Personal Ontology Recommender (PORE) which was created by the National Chung Hsing University of Taiwan and also was being used by the university’s library and Ex Libris’ bX recommender which is available to subscribing institutes were selected from the class of

recommender systems to be analyzed in this paper. Moreover, MyLibrary, which was one of the earliest attempts to provide users personalized information spaces on library portals, was also selected as a case to be explored. Those cases were chosen purposefully for they exemplify different approaches to library personalization and it is my hope that the result of their analysis coupled with the literature review will provide comprehensive answers to the research questions.

2.2. Method of Data Collection

2.2.1. Document Analysis

This study mainly relies upon analysis of primary literature such as original articles, blog posts and web documents as the main source of information. Most of the documents and blog posts were originally written by the respondents. The following table contains the explanation:

Case	**Material analyzed**	**Prepared /written by/available at.**
Huddersfield University Library Opac	Blog posts of library systems manager	David Patton, Huddersfield University Library systems manager • http://www.daveyp.com/blog/archives/1453 • http://www.daveyp.com/blog/archives/1694 • http://www.daveyp.com/blog/archives/1703
	Copac blog	http://copac.ac.uk/innovations/activity-data/ http://copac.ac.uk/innovations/activity-data/?p=453
BiBtip	Article	• (Mönnich and Spiering, 2008) Dr. Michael Mönnich, head of Collection and Cataloging Department, Karlsruhe University Library Marcus Spiering, University IT services. • (Gottwald and Koch, 2011)
	Website	http://www.bibtip.com/en
PORE	Articles	Liao et al (2009) ,(Liao et al, 2010) articles written by researchers at National Chung-Hsing University computer science department and others
Ex Libris bX	Ex Libris Group website, Usability tests	(Ex Libris, 2011), (Thomas et al ,2011) (Ponsford et al ,2011) http://www.exlibrisgroup.com
MyLibrary	Articles	(Ciccone,2005) NC state Library (Shedlock et al, 2010) Galter Health sciences Library

Table 2.1. Source of Data

2.2.2. Unstructured questions

In an effort to get more information on the cases in addition to the examination of respective documents listed above, e-mail communications were used to send unstructured questions to the libraries running the services. The following table shows the correspondences and the time they have been made.

Case	e-mail Address	Sent	Reply
University of Huddersfield Library	AskALibrarian@oclc.org	17 March 2012	19 March 2012
	d.c.pattern@hud.ac.uk (Library system manager)		19 March 2012 (referring to his blog posts)
		23 March 2012	23 March 2012
Bibtip	infodesk@ubka.uni-karlsruhe.de (Karlsruhe University Library)	28 February 2012	14 March 2012 (referring to BibTip GmbH)
	boris.koeberle@bibtip.com (BiBtip GmbH)	24 March 2012	27 March 2012
PORE	janetc@dragon.nchu.edu.tw	22 March 2012	28 March 2012
	javan.nchu@gmail.com	29 March 2012	12 April 2012
bX Ex Libris	h.k.riis@ub.uio.no (University of Oslo)[5] Regina.Lein@ub.uib.no Irene.Eikefjord@ub.uib.no (University of Bergen)	10 May 2012 27 February 2012	14 May 2012 4 June 2012 5 March 2012 9 March 2012
MyLibrary@NCstate	kacollin@ncsu.edu	19 March 2012	19 March 2012 (referring to the article she wrote on MyLibrary@NCstate)

Table 2.2: e-mail correspondences

An interview was made to locally accessible library where they subscribe to Ex Libris' bX recommender system. Beside the above mentioned data collection techniques, the OPACS of Huddersfield University and Karlsruhe University Libraries and also bX recommender on Ex Libris primo were examined.

[5] Unstructured interview was also made

2.3. Method of Data Analysis

As Dey (1993) wrote, the core of qualitative analysis lies in processes of describing phenomena, classifying it, and seeing how concepts interconnect. The intention of this research paper is to thoroughly analyze selected cases, discover themes as they reveal themselves, compare them with related themes in other cases and also the conceptual literature review, and then group them under the research questions they answer. Therefore, constant comparative method was used because this method enables to categorize findings in order to identify main themes and issues that would describe the subject under study (Boeije, 2002).

The constant comparative method was further discussed by Corbin and Strauss (2008) as a process involving three levels of data analysis: Open Coding, Axial Coding and Selective Coding. The word 'coding' is used to refer to the process of data analysis. During the Open coding phase, the researcher compares data and continually asks what is and what is not understood. Systematic analysis of the documents at this stage helps to identify categories, properties and dimensions within or among data. During the second phase, data are pieced together allowing connections between categories. The inductive and deductive thinking process for relating subcategories and categories takes place at this stage. Finally the process of choosing core categories and systematically connecting it with sub categories, validating similarities and relationships then completing categories concludes the data analysis process (Corbin and Strauss, 2008; Kolb, 2012).

Using the constant comparative method as a framework for analysis, the research questions formulated in this research were considered as core categories. In the process of literature review, documents were consulted for themes that would answer the questions and those themes were identified and presented as categories. During the analysis of selected cases, the themes extracted from the cases and the literature reviews were compared. Categories and subcategories were also identified. Finally they were organized under the research questions they answer and relationships and similarities between the categories was shown using the *see* , *cf(compare)* and other references.

Chapter Three

LITERATURE REVIEW

The main task of a university library is helping teachers and students in the teaching, learning and research processes which can be taken as its key success factor. The ability to access knowledge and information, analyze and innovate determines the quality of research results by academic library users and also the future of a society (Lan-Xia, 2010). However, the flood of information to university libraries through print and non-print resources, databases and network of information services poses a challenge for accessing the right information at the right time.

The challenges academic libraries face are part of the challenges the whole academic environment is facing. University libraries are expected to serve diverse group of users from diverse backgrounds. In addition to the subject areas or specialties of users, the list of diversity can go on to include teachers, students, researchers, tech savvy younger users (generation Y), continuing and returning students, etc. who all need information in full text form under their fingertips (Wu and Liu, 2001). Academic libraries may even serve their alumni who prefer timely delivery of filtered and concise information (Helfer, 2002; Smith et al, 2007).

The whole notion of personalization is about helping users to deal with the threat of information explosion and enabling them to access the right information with less time and effort, saving their time for more intellectual activities (Jeevan and Padhi 2006) .Therefore this suggests university librarians to cater to the needs of their various groups of users and be able to provide them with targeted, timely, intelligent and interactive information services (Lan-Xia, 2010).

As it is explained by Brusilovsky et al (2010), personalization and recommender services shouldn't be considered new to academic institutions since they have been manifested in different forms. For instance, we can see them through teacher recommended reading lists,

library course reserves ,and reading rooms and services adapted to groups of users (e.g. Biology freshmen, Masters Students' reading rooms, etc.). In addition to this, we can see recommender services reflected through the services of expert reference librarians putting resources into focus including those which might have been overlooked by a teacher or arrived recently. Moreover, we can also include 'social navigation', where users follow what their pears have read and also the "wear and tear" of books becoming symbols of their importance and high usage. Therefore what personalization requires in modern university libraries is to replicate such past known experiences and devise better ways of utilizing them (Brusilovsky et al, 2010).

Personalization is a service pattern that aims at ensuring different service strategies, contents and functions for different customers. As Qian et al (2007) stated, such service would have the following practical significances if adopted in libraries/digital libraries. First, it brings user at the center which would have high importance for construction of third generation digital libraries which are characterized as being distributed, isomerous and personalized. Second, personalized services ensure sustainable development of a digital library and third, it ensures cultivation of individualization, promotes social development of variety and help save time of customers (Qian et al, 2007). Moreover, it gives a sense of ownership and partnership for users as it provides them with the means to tailor the services according to their knowledge, skills, needs and preferences (Renda and Straccia, 2005).

3.1. Adaptive and Adaptable Personalization

Frias-Martinez et al (2009) classify personalized services according to the three basic services provided by digital libraries namely content personalization, interface personalization and personalization of Information filtering and Information retrieval. They also classify them as automatic or user driven, in other words, adaptive or adaptable (Frias-Martinez et al, 2009; Renda and Straccia, 2005). Users can be the controllers, initiators or the modifiers of the service by providing explicit input of their preferences or it can be an automatic service controlled by the system where the system learns about the user from his/her online interaction and navigational behaviors (Renda and Straccia, 2005). Approaches made to

content personalization via implementation of MyLibrary on academic library websites can be mentioned for the user controlled or adaptable type of personalization (Cohen et al., 2000). As we can see under the chapter presenting selected cases (see chapter 4), recommender systems integrated to library OPACs or search systems can be considered as automatic or adaptive type of personalization.

On the study they made on 60 individuals, Frias-Martinez et al. (2009) found out that more users performed better in the adaptive system and also perceived it more positively. However they also mentioned that they have observed user cognitive styles affecting choices between adaptability and adaptivity. For instance, Field Independent users (users with individualistic behavior) responded more positively to the adaptable version than Field Dependent users (users with more social orientation) whereas Field Dependent users reacted more positively to the adaptive version than Field Independent users (Friaz-Martinez et al, 2009).

From the discussions made above, we can infer that both styles of personalization exhibit their own advantages and limitations that should be taken into consideration. The user controlled system will allow users to enter their preferences and modify the system to suit their needs. But the cost of time and effort to do so might be frustrating. The explicit inputs may help the service to be more tailored to the user needs but the keywords may fail to represent what the user actually wants. On the other hand, the automatic system will not need time and effort from the user side but needs time to learn about the user. As discussed in the following sections, questions remain on the quality and representativeness of the data automatically collected from the user's online activities.

As it is also presented in the following sections, the discussion on choosing either adaptive or adaptable type of personalization spills over to its elements such as usage data, user profiles, user data and also the issue of privacy.

3.2. Usage Data

Usually there are three types of data to be managed on websites: content, structure and log data (Aghabozorgi and Wah, 2009). Taking this to the context of digital libraries or library systems, we can see that such classification holds true. In libraries we have content, metadata and usage data which can include circulation records, log data and data on downloads. Usage data relates information resources to users and provides foundation stone for building users interest knowledge base which is a prerequisite for effective personalized services (see 3.3).

The use of usage data for making various decisions in not uncommon in libraries. Lending figures and re-shelving studies were some of the older techniques for getting insights into the usefulness of the materials and most libraries use circulation or access data to make collection decisions (King, 2009). Usage statistics can be used for making decisions related to staff or facilities and log data can be used to improve design of web pages (king, 2009).

Usage data can be used to evaluate the quality of metadata in representing learning objects and also for personalization and user modeling purposes (Ferran et al, 2007). Moreover, such data, which may also include the history of an individual's interaction in searching information, can be stored in an organized way and be shared to future users which would have similar needs (Ferran et al, 2005).

With the explosive growth of e-journals, many publishers have been providing usage statistics per journals to subscribing institutions. This data was quite instrumental in helping librarians to identify their users' information needs and also used as evidence for funding purposes (King, 2009). Although recording of usage data is common in scholarly information services, its exploitation for the creation of value-added services such as recommender systems remains limited due to concerns regarding user privacy, data validity, and the lack of accepted standards for the representation, sharing and aggregation of usage data (Bolen and Van De Sompel, 2006).

3.3. User Profiles

As libraries work on collection development and access management, they need to give equal emphasis to user interest management (Jeevan, 2008) and develop knowledge base of interests (Qian et al, 2007) which can be referred as the user profile: a representation of user's interests, preferences, needs, goals (Amato and Straccia, 1999) and cognitive styles (Ferran, et al, 2007).

Many definitions of personalization are centered on the concept of user profile. For example, it is defined as "a service provided based on user profile" (Ritz, n.d.) and "a process of gathering and storing information about visitors, analyzing the information and, based on the analysis, delivering the right information to each visitor at the right time" (Chiu, 2001). Fan et al. (2005) regard management of user profiles as the core of personalization technology. Adomavicius and Tuzhilin (2005) attribute success of personalization applications to the knowledge of customers' personal preferences and behavior collected and stored in form of consumer profiles.

Most personalization systems operate on some kind of user profile and the data for the individual user profiles can either be collected explicitly through the user's direct input or implicitly by a push system through agents monitoring user's activity and behavior on an online system (Gauch, et al, 2007; Friaz-Martinez, et al., 2006; Fan et al, 2005). There have been researches on which type of user profile might serve better and, as it is discussed in the above sections, the choice can be a factor of different variables including the cognitive style of a user (Friaz-Martinez, 2009). However most of the researches discussed on this paper seem to tilt towards the implicit user profiles citing the problems explored on explicit user profiles.

One problem of explicit data or keyword-based profiles is the vagueness they might present (e.g. polysemy and synonymy). Such user profiles may fail to accurately represent the user interest (Jiang and Tan, 2009) and face the same vocabulary problems search engines such as Google and Yahoo! are facing(Fan et al,2005) . To put in in a nutshell, keyword based user profiles fail to capture the semantics of user interests (Degemis et al, 2007).

There have been solutions forwarded to solve the vocabulary problem of explicit user profiles. One solution forwarded by Degemis et al (2005) suggests making them semantic or concept-based because such profiling would help in forming concept hierarchies and relationships , enabling a user to get his/her materials though he/she didn't use the exact keywords for retrieving the material.

Though the idea of semantic user profiles is getting traction in the area of information retrieval, Jiang and Tan (2009) said its application in personalized services is minimal. They presented a semantic based user model called User Ontology for supporting personalized applications in the semantic web. They said this model utilizes concepts, taxonomic relations, and non-taxonomic relations in a given domain ontology to capture the users' interests. They also added that the proposed model has been integrated into a semantic search engine called OntoSearch to provide personalized search, applied to document retrieval in the ACM digital library and the Google Directory and showed encouraging results. Liang and Ku (2008) proposed a semantic-expansion approach to build user profiles and content recommendations. According to this method, user profiles are constructed first by extracting concepts (keywords) from the users' reading history. Those concepts are treated as representations of user interest and will be expanded using a semantic expansion module which is consisted of a library of semantic trees and a set of spreading rules. According to them, this approach is better than the traditional keywords approach in capturing user interests but the major concern would be how to build comprehensive and useful semantic-expansion networks to cover major concepts and their relationships (Liang and Ku ,2008).

The other problem of explicit-input-based user profiles is a possible reluctance of users to evaluate an information artifact and provide feedbacks (Perugini and Gonçalves, 2002). These amounts to data sparsity and other related problems discussed more under the section dealing with collaborative recommender systems (see 3.6.2)

Based on the systematic comparative study they made on profile generation methods, Fan et al (2005) concluded that implicit profiles are superior to explicit profiles. Moshbar et al (2000) also share this idea by explaining that, unlike the explicit ones, implicit profiles will not decay through time, avoiding dependence on registration based personal preferences and

potentially subjective user ratings. Studies suggest that user's behavior can be identified from their history of interaction with incoming items as the result of knowledge acquisition process and there is direct correlation between an item's relevance to a user and the amount of time the user spends on it (Morita and Shinoda, 1994 as cited in Hanani et al, 2001). But, a user might retrieve the item and be interrupted by other tasks such as phone calls therefore this time-spent-per-item measurement makes implicit inputs open to bias (Hanani et al, 2001). Therefore such inputs may have to be complemented by some kind of user interrogation such as relevance feedbacks (Hanani et al, 2001). But, it would be worth noting that users may perform either exploratory - just to check what materials are available, or goal oriented navigation where a user is looking for a resource (Ferran et al, 2005). Such fact adds complicacy to judge whether user's search activities reflect his/her interests.

Aside from the discussion of choice between implicit or explicit input driven user profiles, there are some other works showing the importance of adding some more elements to user profiles. Friaz-Martinez (2006) stated user profiles in digital libraries can have eight potential dimensions such as device, context, history, interest, goal, domain expertise and human factors in order to represent a user (Friaz-Martinez, 2006). Stewart et al (2004) added that most works on user models are limited to cognitive pattern dimension such as user interests, skills etc. which can present a major obstacle for realization of personalized services in digital libraries. Therefore, they recommended the addition of context to the user model. According to them, context may include user's position in his/her organization, relationships and community roles, etc. Illustrating this proposal, they presented an extended user model which has three components: Domain Ontology, Resource Network and Personalized Web Context. The domain ontology describes all entities and relationships in the domain. The domain can be, for instance, the scientific community which would include people, events and publications. The Resource Network (RN) is set of information models with nods representing resources in the domain and the type of relationship existing between them. The personal web context (PWC) represents relationships and type of relationships between a user and an entity. (Stewart et al, 2004).

Nika et al (2011) raised the need of adding interoperability in terms of context and structure to the user profiles. The rationale behind this, according to them, is that users are interacting

with different digital library systems and subsequently their user information would be scattered across those systems. This would call not only for digital library interoperability but also for user data interoperability to reconcile user information residing in them and achieve cross-digital library personalization (Nika et al, 2010). Ferran et al (2005) also adds that user profiles can be modeled by ontologies and when a digital library user leaves one service to connect to the other one, the user profile too can be transferred through the appropriate semantic web technologies.

All the discussions made above on user profiles, the choice between implicit and explicit input, application of ontology to create better user models, need of adding context, interoperability and other elements to a user model, are about making the user model rich and more representative of the intended user because that is what determines the effectiveness of personalization (Gauch et al, 2007). But what does that mean to user privacy?

3.4. Privacy Issues

Libraries have a long history of protecting their users' privacy which is now being challenged by government regulations like the USA Patriot Act, which allows law enforcements greater access to library records (Fifarek, 2002) and technologies like web 2.0 tools which not only add highly interactive and information sharing environment, but also may create loopholes for compromising user privacy (Magi, 2010). The potential of technology in collecting information on any online behavior of users seems to have alerted academic librarians to take more measures in protecting their patrons' privacy. As Fifarek (2002) discussed, laws that protected privacy of user circulation records for print collection don't say enough about defending user privacy in the online environment. Therefore she advices librarians to delete old log files or, if they need to keep them, to clean traces that might lead to an identifiable individual ,or contact software vendors for tools anonymizing the system logs (Fifarek, 2002).

The firm stand and concern libraries have on privacy seemed to have stretched beyond their traditional turf. For example, Magi (2010) said features such as EBSCO's 'MyEBSCO', Elsevier's 'My Settings' and others available on vendor databases pose threat to privacy. She argues that

despite of the fact that libraries have done well in protecting privacy when users search online catalogs, check out materials or ask reference questions, they have not addressed the potential threat posed by vendors of web-based information resources. On the research she made on twenty vendors' privacy policies, she pointed out that their policies fail to express commitments to many of the standards outlined by the profession of librarianship and information technology industry regarding management and protection of user information. With increasing economic values of personal data especially in market research and companies investing millions of dollars on personal-information-rich social networks, Magi(2010) cautions that some database vendors might be motivated to sell user data . Therefore she urges librarians to examine privacy policies of their vendors and advocate for protection of user privacy.

The discussions like the above put libraries at an interesting position with regard to personalization. On one hand, improving user experience requires collecting information about the users (Awad and Krishnan, 2006). But on the other hand, however, libraries are required to operate under stringent privacy practices and requirements. For instance, many digital library systems maintained by the U.S. Government are not allowed to collect data about individual's download behavior (Bollen et al, 2007). In Europe too, the culture of privacy is so strong that any attempt to profile users may illicit negative responses from customers (Tv Genius, 2011). However, it could be worth noting that applications such as Google iTunes and Amazon.com run successful recommendation services which involve utilization of user and usage data(MacManus, 2009).

As book retailers such as Amazon.com retain their user's data with a policy framework to protect privacy and provide recommendation services , librarians follow very strict privacy policies (Van Ullen and Germain, 2002) which includes breaking the link between a user and a resource once the material is returned to the library (Lynch, 2001). Yet personalization is recommended as a good solution to the glut of information produced by the digital media and this seemingly presents a dilemma.

As the need for personalization increases, so does the need for privacy. These two seem to oppose each other but need balancing (Gauch et al, 2007). There have been technical and

less- technical suggestions forwarded to strike the balance. The first solution is to make the process transparent and let users know that their actions are logged (Ferran et al, 2005). They should also be assured that their data will solely be used for recommendation/personalization purposes (Ferran, et. Al, 2005). This entails a formulation of a thought through policy framework for protecting privacy.

Adding on the issue of thrust, Almazro et al (2010) suggests that, as it is practiced in any system that uses data mining techniques, users must trust the recommenders to protect their privacy properly and the recommender systems in turn must be built more secure to protect external attacks by malwares or external agents that would compromise the security of user data. Neuhold et al (2003) mention W3C's Platform for Privacy Preferences (P3P) standard as another effort to protecting user data. Using this standard, a user can specify his privacy preferences, defining the purposes his personal data can be used for. The content provider in turn commits to the privacy policy stating the intended use of the data (Neuhold, 2003).

Another approach towards protecting privacy while benefiting from personalization is presented by Jing-Sen et.al (2009) in form of APIRS[6]. This is a system where personalized service runs on client side and retrieval service runs on the server side. The client side is comprised of a user interface module, a personalization processor, a user characteristics database, and a login signer. The server side is comprised of an index manager, a group of index servers, and a login server .When a user logs in to this system, the server's task would be to authenticate the user and perform searches. The user will remain anonymous but valid user to the server. The client side hosts the user interest models and performs analysis on the search results to generate recommendations (Jing-Sen, 2009). Such server-client design would make the personalization function under the control of the user.

Those discussions made above show the need for privacy is legitimate but cannot be an obstacle for realizing personalized services. Standards, laws and technical breakthroughs as discussed above show that it is possible to strike the balance between the needs of personalization and privacy.

[6] It is not clear for what the acronym stands for.

3.5. Personalization Vs information literacy

Beside the issue of privacy, information literacy seems the other area that makes some skeptic of personalization. For instance, Badke (2012) raises the need for looking at the issue from two perspectives: the need to customize tools to best meet our information needs or the intention of vendors to target their goods or services to our preferences. He points dependence on the past history of the user for predicting future preferences as a failure of personalization and argues that it fails to serve the needs of a researcher when he picks up a new research topic he has never dealt with before. Adding more on this issue of "bias to the past", MacManus (2009) reflected on the failure of recommender systems in showing "new things" and their difficulties in dealing with changing data and changing user needs. Therefore Badke (2012) suggests teaching search skills is better than letting users to be locked up into machinations of search history or allow algorithms to disable our skills of information search and make us preys for advertisements.

However, one of the most pronounced challenges of the information age is "locating relevant information in a haystack that is growing rapidly" where personalization techniques can be regarded as 'assistants' to information search with minimal user effort involved (Renda and Straccia, 2005). The user finds an article and the recommender system finds more of them (Ex Libris group, 2011). In this sense, we can see recommender systems as valuable extensions of users own search results. In addition to this, we can understand the goal of personalization as making libraries person-oriented (Qian et. al, 2007) which helps libraries be closer to their individual user needs. Technology has been helping people many ways in life and if it can help them in libraries by delivering tailored and proactive services, it would be welcomed.

3.6. Types of Personalized Services

Efforts to provide personalized information services using information technology have been existent since the introduction of Selective Dissemination of information (SDI) by Hanspeter Luhn of IBM Laboratories in the 1950's(Hensley 1963). This concept was seen as the reverse of

information retrieval where documents find users instead of users finding documents (Hensley 1963 ; Foltz and Dumais, 1992). The SDI was a proactive information service which perhaps can be regarded as the earliest attempt of libraries to provide personalized information services using information technology (Foltz and Dumais, 1992).

As explained by Hensley (1963), SDI was a system that lets users register with keywords representing their interest and that data was organized and stored in the system in the form of user profiles. It involved the use of computers to process the data. When a new publication is added, the system compares it against the user profiles and notifications for new item's arrival will be sent to the users whose profiles match with the content of the document. The SDI was also called "Current Awareness" since the objective was to make users aware whenever a new material of their potential interest arrives (Hensley 1963). As we can see from the discussions in the following sections, this explicit-inputs-based service has evolved through time to include tools such as e-mails, RSS feeds, recommender services and personalized information spaces.

3.6.1. E-mail and RSS .

As the internet grew in size so did the volume of information and also the need to control information. RSS was found to be an answer (Anderson 2006). There seems no clear-cut agreement as to what the acronym RSS stands for. It is been described by different people as "Rich Site Summary", "RDF Site Summary", or "Really Simple Syndication" (Estabrook and Rothma 2007) where 'Really Simple Syndication' seems the favored expression in the field of Library and Information Science (Sauers 2006).

According to Estabrook and Rothma (2007), RSS was seen as a better way of providing current awareness service or selective dissemination of information than e-mail, which has been a tool for such task for many years. the reasons as listed by the aforementioned authors were: RSS doesn't clog e-mail accounts, there is no risk of spam with it and there is no risk of missing an update in the shuffle of other messages (as it is with e-mails), anonymity is maintained as subscription to RSS feeds doesn't require one to submit his or her personal information, and

items can be accessed en mass as opposed to e-mail which requires opening the messages one by one. Adding more to the advantages of RSS feeds, Sauers (2006) puts them as solutions for a busy person who may not find time to check blogs. RSS gives that person a chance to subscribe to the website or blog then the computer takes care of identifying new developments and notify the person whenever something new happens. In a nutshell, RSS saves time of the user by consolidating information from different resources into one place. The user will have total control of what he/she want to see.

The RSS tool is increasingly popular in libraries as it can be judged from the orange RSS buttons on different library websites. But there is a great deal of shortage in literature to see how this tool is really being utilized. Citing this concern and adding a contribution from her side, Nelson (2008) wrote about her experience in a small special library in Canada. The library was providing current awareness service using e-mail alerts. Users appreciated the service but the problem was that the alerts were clogging their e-mail boxes and there was a risk for some of the alerts being left unread. In addition to that, it was clear that using e-mail alerts as current awareness service faced a problem whenever a staff re-assignment or turnover happens. The RSS feeds were the next mechanisms the librarian opted for. The problem was that users don't know how RSS works so it was necessary to give them orientations. After some time she found out that there were still some preferences for the e-mail alerts as there were some people that forget something that doesn't come up front. Then she added weekly e-mail digests. Finally she conceded that introducing such services takes time and energy and would be unrealistic to expect all staff to make use of the system. However, a significant number of staff must use it to make the effort worthwhile (Neilson 2008).

The problem with RSS feeds comes when there are too much feeds a user has to filter. Therefore there are works being done to tweak RSS and make it easier to use. One of such efforts includes designing RSS to acquire user feedbacks implicitly, avoiding the labor of explicit adjustments (Samper et al, 2007).

3.6.2. Recommender Services

Following the spread of internet from early to mid-1990's, recommender systems were introduced to various commercial applications to help users address information overload (Konstan, 2012; Park, et al, 2012) and to retain customers (Perugini and Gonçalves, 2002). The idea of incorporating user feedbacks in form of ratings and using the ratings of "like-minded readers" to recommend resources has been existent in music, movie and UseNet news applications (Perugini and Gonçalves, 2002). This type of collaborative filtering expanded gradually to include content based and knowledge based applications (Konstan, 2012).

Recommender Systems can be defined as software tools and techniques providing suggestions for items of potential interest to a user (Ricci et al, 2011) . They are the most popular personalization approaches used in information and content management consisting of three major components: 1) background data - the information that the system has before the recommendation begins; 2) input data – information the user must supply to the system in order for a recommendation to be made; and 3) an algorithm that combines the background and input data to arrive at a suggestion (Neuhold et al, 2003; Ricci et al, 2011).

Recommender systems are considered as techniques of information filtering and based on the techniques they follow they are classified as collaborative, content-based, hybrid (of collaborative and content based) and knowledge based (Jannach, 2011). The content-based and collaborative types are those recommenders which are the most used and most treated in literature and this section limits itself to discussing them.

3.6.2.1. Collaborative Recommenders

The notion of collaborative recommenders is that if a group of users shared the same interests in the past, they would also have similar tastes in the future (Jannach et al, 2011). Therefore the primary interest of techniques employed by such type of recommenders is to store and process data of user preferences. As illustrated by Jannach et al (2011), if two users A and B have similar history of, for instance, buying books, it is considered as the two are implicitly collaborating with one another. Therefore when user A buys a new book, the

recommender system would suggest this book also to user B. As expounded more by Degemis et al (2007), Traditional collaborative recommendations work by calculating similarity value between a current user and other users by considering set of ratings given on same items. Users with most similar styles of rating will be considered as neighbors and based on that similarity; collaborative algorithms compute recommendations for the current user. Collaborative filtering is also considered as a form of Social Navigation as it is based on "collective wisdom" of people to guide future users (Brusilovsky et al, 2010).

Collaborative recommenders have been successful in e-commerce and are getting traction with libraries where they have a potential to improve information (see table 4.2, for example). However they have exhibited problems which prompted researches in making them better.

One issue with collaborative recommenders as discussed by Im and Harris (2007) is the problem of bias they would surface when they use explicit user evaluations. Users might have different intentions and information needs when they make those evaluations. This possible mix of intentions would create bias which could decrease the accuracy of the recommendations (Im and Harris, 2007).

Discussing on user ratings, which constitute a type of explicit inputs being used by collaborative recommenders, would demonstrate the problems such inputs carry along. As illustrated by Geisler et al (2001), Ratings might be exaggerated or wrong. For instance, if we ask a user to rerate an item, he may rate it differently than the way he did before. Moreover, we cannot be sure of getting the feedbacks as users might be reluctant to rate an item if they are not clear with the benefits they get from the process (Perugini and Gonçalves, 2002). Some users might rate items while the others don't causing sparsity, one of the main problems that plagued collaborative recommender systems (Vellino ,2010).

There are suggestions being given to solve the problems related to sparsity. Some suggest use of compensation system for rewarding those who rate items, the others propose looking at the user behavior and derive ratings implicitly and also the use of dynamic agents to automatically rate items (Almazro et al, 2010). Vellino (2010) suggests it is better for collaborative recommenders to be based solely on usage data. Therefore implicit feedback

systems such as views, clicks or queries have become more popular in the current collaborative filtering systems (Guy, et al, 2009).

The other problem that can be raised in connection to the use of usage data just discussed above is a possible lack or shortage of data (MacManus, 2009). To get good recommendations, more users are needed. The more users use the system, the more usage data will be generated and the more robust the recommender system will be (MacManus, 2009). If the usage data per an item is so little then collaborative filtering might not work causing a situation pronounced as 'cold start' problem (). Combining content and usage mining or creating the hybrid of collaborative and content-based recommenders was provided as part of the solution (Moshbar et al, 2000).

Adding more to the problems of collaborative recommenders, Almazro et al (2010) caution that there could be a chance of missing lots of good items from recommendations because no one has rated them. A new item would not be recommended until it is well used and rated by a sizable number of users (Adomavicius & Tuzhilin, 2005).

Moshbar et al (2000) raised scalability as a problem in collaborative recommender systems. According to them, scaling collaborative filtering to larger number of items while maintaining reasonable prediction, performance and accuracy is difficult. They suggested clustering user records with similar characteristics and focusing the search for the nearest neighbor only in the matching clusters would be a solution.

Im and Harris (2007) say that the success of collaborative recommendations depend on users' search mode. Citing the experiment they made on two product domains, they explained that collaborative recommendations work better when users perform specific searches rather than general searches. Therefore they explained that such recommendations should handle user's mode of search (Im and Harris, 2007). This leads us back to the discussion of context made under section 3.3.

3.6.2.2. Content-based recommenders

Content based recommender systems work by matching item-features to user profiles (Ricci, 2011). They depend on the active user's known preferences to identify and recommend items with features similar to the items the user has liked in the past (Lee et al, 2009). Unlike collaborative recommenders which are based on user opinions and similarities of opinions by user groups, content-based recommenders approach the issue of recommendation as a search for related items, attempting to recommend items similar to those a user has liked in the past (Almazro, 2010, Degemis et al, 2007).

Adomavicius & Tuzhilin (2005) point out overspecialization as a major problem of content based recommender systems. According to them, such systems fail to recommend anything different than what a user has seen before. As Perugini and Gonçalves (2002) put it, in those types of recommenders, frequently purchased items (such as banana in grocery market) will always be recommended and products like cars that are seldom bought would face risks of not being recommended at all. Adomavicius & Tuzhilin (2005) also add that some content based recommenders might ignore recommending items which are too similar to what the user has seen before considering them as redundant. The other problem is the "new user" problem meaning a new user will not get any recommendation until he/she makes many ratings (Adomavicius & Tuzhilin, 2005; Perugini and Gonçalves, 2002).

3.6.2.3. Challenges of Recommender systems

Konstan (2012) mentions risks to individual privacy and structure of social relationships as hidden dangers of recommender systems. If not deployed thoughtfully, these systems would put walls between different communities and obstruct communications between people of different communities. They would also cause loss of individual privacy if privacy aware recommendation techniques are not used (Konstan, 2012).

As Geisler et al (2001) pointed out, there are different sources of data that can potentially be used to base recommendations upon. For instance, resource descriptions, resource usage data, ratings, are some of the categories. With all problems related to data sparsity, scarcity

and quality as discussed in the previous sections, they explained the choice of a dataset to base recommendations upon poses a challenge. As they explained the problem further, some users might rate items while the others don't, some would have profiles while the others won't. We can only be sure of the existence of basic resource descriptive data rather than reviews, ratings or any other type of data that has to be explicitly recorded by a user. Such data might be noisy, corrupted or just plainly wrong (Konstan, 2012). Ratings have been the most used explicit user inputs but the challenge ahead includes utilization of tags, reviews, tweets, facebook updates and other types of user generated content to improve the database of recommender systems while preserving privacy (Konstan, 2012).

Therefore as explained by Geisler et al (2011), choosing a dataset to base recommendations upon would be challenging. If we use only one set of data, for instance, download history, then there is a potential of ignoring other rich data sources that would be good for generating recommendations. On the other hand, if we design the recommender systems to get data from any available source, then the variable quality of resources would threaten the usefulness of the recommendations (Geisler et al, 2001). The problems with recommender systems are not limited those discussed above but there are also some other issues such as relevance of results, duplication of recommendations, ease of use, appearance, etc. (Reisinger, 2009) which would require further research and development in the field.

One of the main problems with the traditional recommender systems as discussed in the above two sections is the cold start or a new user problem. As explained by Middleton et al (2002), the cold start problem happens when no initial information is available early to base upon the recommendations. As discussed in section 3.3 and also as explained by Middleton et al (2002), the application of semantic web and ontology is believed to help in solving the problem because they can provide valuable domain knowledge and user information .

3.6.2.4. Not Personalized Enough?

As it is discussed earlier, the existing collaborative recommender systems help users to deal with information explosion. But they are not personalized enough as they work on collective assumptions or group preferences (Pera and Ng, 2011). They recommend products or services

to customers based on what other customers of similar tastes have said about the products. The recommendations are independent of the customer, so each customer in the group gets the same recommendations. Shaffer et al (1999) classified them as non-personalized recommender systems and refers to them as 'ephemeral' because they don't recognize the customer from one session to the next. Such recommender systems are common in physical stores (Shaffer et al, 1999).

Adomavicius & Tuzhilin (2005) recommenders will yield better results if they are made to be context aware. According to them, the current generation of recommender systems operate in two dimensions: user and item information. But they fail to add context which may include time of the year, the person's background, company, etc. which might be crucial in some areas. Therefore we can say this third dimension can make recommendations more personalized.

So where can be the source of this 'context'? Belkin and Zhang (2004) believe that we can get it out of the implicit data recorded as users interact with information systems. Pera and Ng (2011) proposed the use of social networks to make recommendations more personalized. They developed a Personalized Recommender that relies on Friendships (PReF) established by users of social website such as LibraryThing. This system makes use of the user's friends, personal catalogs of friends and their ratings to come up with recommendations (Pera and Ng, 2011). Elaborating on the need for including social networking information to recommender techniques, Liu and Lee (2009) explained that those recommendation techniques fail to distinguish friends in a neighborhood from strangers who have similar tastes. Therefore they recommended adoption of a hybrid approach for utilizing social network information with collaborative filtering methodologies that have been used by the recommenders (Liu and Lee, 2009). Inclusion of social networks helps not only in exploiting the huge user data they have but also helps in recommending people who may have the same interest or inclination (Guy, et al, 2009).

User profiles are the other obvious potential sources for user context. Though the recommender techniques discussed so far help in tailoring information to a potential right

user in some way, the ultimate prerequisite for ensuring personalized services lies on user profiles (Amato and Straccia, 1999).

3.6.3. Personalized Information Spaces: Academic Library Portals

Large academic libraries host digital contents residing in different systems, accessed through different interfaces and authentication procedures which apparently frustrate users. Therefore the need for a library portal that gives integrated access to all resources through a single interface and authentication process becomes apparent. Cecelia & Yang (2005) presented the university of Singapore library portal as an example for such endeavor. According to them, the portal was developed using enterprise portal technology to deliver personalization and single sign-in capabilities. Beside the advantage such system provides in terms of seamless access to all digital contents of the library anywhere with a single log on, they added that the system enables users have their own profiles and create their own personalized research and information environment based on the available library resources.

Another example for integrating various information resources in a university and meet user information needs in simple and personalized way is presented by Yin and Peng (2009). According to them, Shanghai Jiao Tong University Library uses MetaLib and SFX to integrate their resources, provide single access point to all of their resources, and make use of personalization features on MetaLib. Those features as listed by Yin and Peng (2009) are:

- eShelf. Users can mark records from their search results lists and store them in their personal eShelf for future use.
- My databases. Users can create lists of selected databases, assign them a name and use them for cross searching.
- My e-Journals. Users can create lists of selected journals, and browse and access them or obtain services through SFX.
- History. This can enable users to store a search executed in a previous session, in order to run it in the future or to create an alert.

- Alert. An alert is a query that searches the specified set of databases automatically, at an interval of the user's choice. An alert will notify the user by email when new records that match the search criteria are added to the specified databases.
- Preferences. This can enable users to personalize their own interface environment, such as the language, the number of results per page, etc.

Nicholes and Mellinger (2007) however advise that libraries need to make a good test to ensure that personalization matches information needs before they commit significant investment on creating portals and risk a poor rate of return. On a study they made on 88 undergraduates at the Oregon State University, they found out only 14 percent of the respondents said they would personalize the university library's website. Upper division of students were found to have more positive view to personalization than the lower division undergraduates which are less experienced to the library and its website and find it as confusing and time-taking. (Nicholes and Mellinger, 2007).

3.7. Trends in Academic Libraries

3.7.1. Personalizing Information Spaces

The future of academic libraries is entangled with the developments happening in the field of education. With the continual transformation of the educational landscape as being resource based, student-centered and collaborative, the libraries need to reflect those changes. As Bryant et al (2009) have put it, those changes are already being reflected on the physical library which include transformation of the library as learning resource center by repurposing the physical space to Include reading rooms with computers for querying of electronic and audiovisual materials collection, collaborative study spaces, presentation facilities, laptops, whiteboards, data projectors, areas for group study, and other facilities (Bryant et al, 2009).

With the expansion of virtual campuses and virtual libraries available 24/7, and also increasing volume of electronic documents, libraries would need to think about replicating their experience of repurposing the physical spaces to the virtual spaces (Savin-Baden, 2007). This may take a form of personalized Information spaces where users may organize information

according to their own interest, with advanced features to enable collaborative work among the users (Renda and Straccia ,2005). This would enable the system to generate recommendations by exploring relationships between users and user communities and preference pattern of users (Avalanchi et al, 2007).

3.7.2. Enhancing OPACs

Can OPACs remain relevant in the future academic libraries? Wallis and Kroski (2009) say OPACs need to be kept up with web 2.0 technologies in order to remain relevant for the next generation of users. Mi and Weng (2008) urge academic libraries to aim at designing self-sufficient, twenty first century online catalogs that fit the web 2.0 model so as to make users comfortable and confident using their library OPACs. The authors further caution that the future of academic libraries depends on the effectiveness of their OPACs as they cannot afford to be seen irrelevant to the information seeking world. Wallis and Kroski (2009) compared OPAC with the Amazon's spell check , "you might like this "service, ''did you like this" service or to the Youtube's on media on demand and other features such as social tagging and concluded that OPAC, which once was an inspiration for these technologies, is now left behind (Wallis and Kroski, 2009). They further added that "Make it more like Google" is a catchphrase often heard from students at academic libraries when they are asked on how library catalogs can be improved.

Wallis and Kroski (2009) further added that the closeness of the new generation of users to technology has raised the bar for libraries and the likes of iTune, Amazon and Google are now seen as standards to measure the performance of libraries (Wallis and Kroski, 2009). MacManus (2009) called Amazon "king of Recommendations" and mentioned that its recommendations are based on "individual behavior, plus either the item itself or behavior of other people on Amazon." He also mentioned Google's personalized recommendation effort explaining the recommendations might be item based as "did you mean..." feature or social recommendation for example, based on who is linked to the webpage. As it is discussed at the start of this thesis under section 1.2. the use of such recommenders in library OPACs is minimal.

There have been efforts done by academic libraries both in terms of providing personalized information environments and enhancement of their OPACs. The following chapter presents the cases which will be followed by analysis with the ultimate aim of answering the research questions of this study.

Chapter Four

PRESENTATION OF SELECTED CASES

This section presents selected cases for personalized service application in academic libraries. They are used to portray in what areas of library activities those functionalities may be applied, what they need to succeed and what problems may lie ahead. The cases were primarily discovered in the course of literature review but later enriched with feedbacks obtained through communication with the libraries and consultation of their blogs, websites, and revision of primary literature (the whole process is explained under section 2.2). The cases represent different approaches to library personalization and it is the hope of the author that their presentation here and subsequent analysis and discussion in the following chapters would provide insights and perspectives as to how the issue can be approached in academic libraries. Two types of personalized information services are identified in this paper: recommender systems and personalized information spaces. The first four cases fall under the first group whereas the fifth one represents the latter.

4.1. The OPAC of Huddersfield University, UK[7]

The university of Huddersfield library has been accumulating library circulation data since 1996. When the amount of transactions logged reached over three million, the library decided to try creating some useful services for students and in 2005 they started working on adding Amazon- inspired "those who borrowed this also borrowed…"functionality to the library OPAC. The service went fully live in January 2006. The library claims that the "people borrowed this also borrowed…" suggestions were popular and received hits with a peak of 5,229 clicks in a single month. The library also found broad similarity between the borrowing graph and the number of clicks on the recommendations per month, implying that the students might have been helped by the recommendations to check out related books when

[7] The information for this section was obtained through consultation of blogs maintained by the library systems manager David Patten, email communication with him and consultation of the OPAC. The methods of data collection used for this and the following cases is detailed under section 2.2.

they couldn't find the books they want to borrow. The library also noted that there was an increase in the circulation of books after the new functionality was introduced.

This recommender service was designed to give three initial recommendations for a hit, with option to see more or everything. By clicking the "more" option, the user can get three more recommendations. Clicking on "everything" option gives list of all available suggestions for a given item. It was obvious that the large volume of circulation data the library has been accumulating through years has played to the advantage of this service.

The issue of maintaining user privacy was also addressed. Therefore the data was aggregated and anonymised. As the techniques used by the recommender system are data mining techniques, the focus was on identifying sizeable groups of users who show the same behavior rather than looking for unique combinations of borrowing that might relate to an individual. Therefore this process helped in striking the balance between protecting privacy and allowing usage data to provide recommender services.

While exploring the OPAC, we can see that recommendations are not necessarily available for every collection the library has. It appears that newer materials or those items which are not borrowed by users would have no chance of appearing among recommended lists. It is also possible to discover a different type of "Other editions and related works…" recommendations available for some hits displaying other works by the same author or different editions of the same book.

This OPAC with a recommender feature has been running for more than six years but it has been centered on the physical collection of the library. Now the library is pondering on extending the service to include its electronic collection. Currently there an experiment going on introducing a "those who looked at this thing also looked at…" functionality utilizing the "e-stuff" data collected from the university link resolver, library management systems and EZProxy logs.

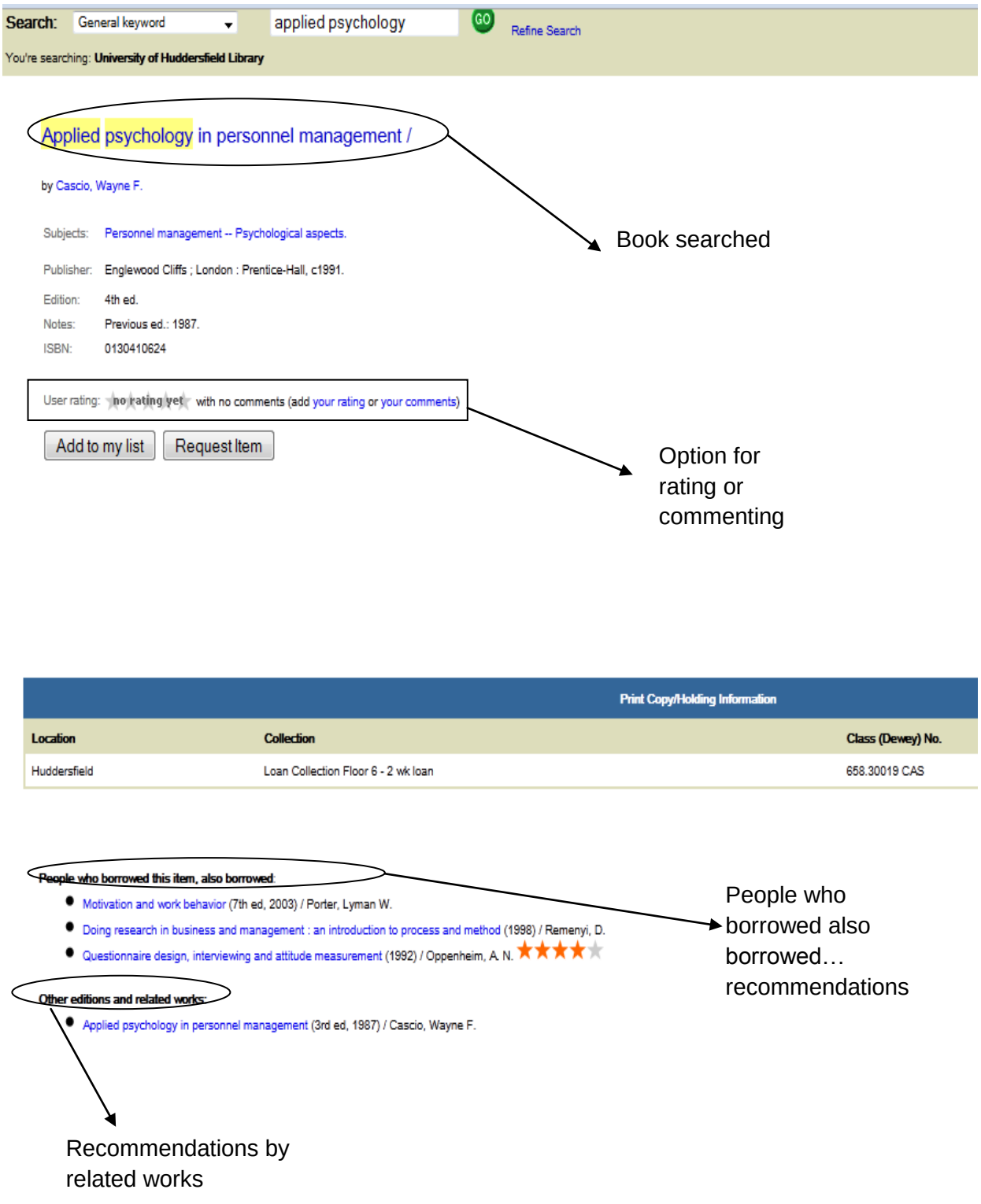

Figure 4.1.OPAC: Huddersfield University Library

The Huddersfield university library have released the major portion of their usage data (circulation data and recommender data) under Open Data Commons/CCO license to the Copac[8] Activity Data Project (SALT 2) which is funded by JISC (Joint Information Systems Committee) of UK[9] . As it can be learnt from SALT 2 blog (copac, 2012), JISC funding illustrates the importance given for aggregating and sharing library circulation data to support recommender functionalities at local and national levels. The aim of the project is to aggregate and normalize data from libraries and share the data to support recommender functionalities at local and national levels. Currently it has partnership with universities including the University of Manchester which has accumulated circulation data for more than ten years, Cambridge University Library, Lincoln University Library, Sussex University Library and University of Huddersfield Library.

There has been a test made on the usability of SALT recommender system which was based on static circulation data from the University of Manchester (Copac,2012). The users which were post graduate students were asked to evaluate the recommender. It was found that there is a unanimous support for the system and the majority view it as important. But in the process it was found that there were recommendations which didn't seem important for the searchers. The lists of suggestions included old and outdated books and also books which users know before and don't suite for their 'niche' research purposes (Copac, 2011). Such weaknesses call for improvements.

4.2. Personal Ontology Recommender (PORE) at National Chung Hsing University of Taiwan

The Personal Ontology Recommender (PORE) is a recommender system developed and used by the National Chung Hsing University of Taiwan. According to the library, the system is not working currently as the library replaced its automation system with a new one last year. They said that it is suspended because of compatibility issue with the new system but will resume functioning after they finish the renewing process.

[8] Copac is a national catalog in the UK bringing together the catalogues of over 70 major UK and Irish libraries

[9] http://copac.ac.uk/innovations/activity-data/?p=453

As presented by Liao et. al (2009), PORE works by building personal ontologies on the patrons borrowing records. User profiles contain keywords of user interests. The researchers have put two main benefits of using borrowing records for building the recommender systems. The first advantage they mentioned was decreasing the probability of faulty recommendations, for instance, like recommending a book about computer virus for a person who is interested on biological virus. The other advantage of borrowing records is that they reflect the changing user interests. This model was first built for the Chinese collection and was based on the Classification Scheme for Chinese Libraries (CCL) as reference ontology and was later enhanced to include English collections, using the Dewey Decimal classification (DDC) as a reference ontology (Liao et al, 2010). Individual users will have their own unique personal ontologies which are built from keywords of their interest stored in their profiles. Those keywords are extracted from their loan records.

Explaining the keyword extraction process, Liao et al (2009) explained that The Chinese word segmentation system (CKIP) developed by Academia Sinica in Taiwan was adopted to extract the primary keywords from book titles and other related information first, and an algorithm was developed to calculate the distinctness levels of the keywords. Explaining about keyword extraction for the English collection, Liao et al (2010), explain that A tool called "the part-of-speech (POS) tagger for English", developed by Tsujii Laboratory at the University of Tokyo is used to identify each word as a noun, verb, adjective, or adverb, etc from a library item record that includes author, title, etc. Then personal ontologies will be built based on the keywords.

The enhanced version of PORE has also included collaborative filtering technique that identifies users with similar personal ontology and then recommends English collections from common interested topics (Liao et al, 2010)

To get additional information on PORE in addition to the cited documents above, the library was contacted via e-mail to get answers for two questions. The first question was about who creates or updates the user profiles used by PORE and how the privacy issue is addressed. According to the explanation from the library's division of information systems, there are already user profiles on the Integrated Library System (ILS) which were originally created by

librarians but which can also be modified by users through their library account. Those user profiles are exported to PORE. The exported user information includes only users' system ID, name, gender, unit code, and identity code and loan data. PORE uses ILS's API to carry out the authentication of user status. The user information will be updated to PORE every month.

The second question was how long would it take for PORE to understand a new user, his/her interests and start recommending resources. In other words, how it handles a possible 'cold start' problem. The answer was that, since students are the main users of the library, user preferences of each academic unit/department are analyzed in advance and respectively set as the "Default Personal Ontology." When a new user first login to PORE without any personal Ontology, the system will start the recommendation based on the "Default Personal Ontology" of his/her own department.

4.3. BibTip

BibTip was developed in partnership between Karlsruhe University Library in Germany and the Institute for Information Services and Electronic Markets (Mönnich & Spiering, 2008). The institute was responsible for developing the algorithms and the scientific basis and the library was responsible for integrating the system to the university catalog, collecting statistical data and developing and implementing BibTip as service. The recommendation system uses implicit feedback and it is based on the behavioral patterns of users interacting with the catalog. The architecture of the system involves three software agents: the OPAC observation agent that monitors selection to titles during OPAC sessions, the Observation Aggregation agent that aggregates the data and does statistical computations and lastly, the recommendation agent presents list of recommendations for the user. For privacy reason, anonymity for the processed data was secured through the use of session IDs (instead of log in ids) and identification numbers given for each session (Mönnich & Spiering, 2008).

According to the company website[10], BibTip is now being used by 47 universities and colleges and 8 national, state and city libraries.

[10] www.bibtip.com

Unlike the previous three cases, BibTip at Karlsluhe University Library don't use circulation data. The reason as given by Mönnich & Spiering (2008) was that circulation data omits reference books and other literature which are not allowed to borrow. The other reason was that catalog search data is bigger than circulation data so it ensures a quicker way to get the critical mass of data the recommender needs to start operating. Moreover, if recommendations are calculated by circulation, recommendations will be influenced by the availability of the material and, a book that has only one copy but needed much by users may receive low ranking than the one which is less interesting but has multiple copies (Mönnich & Spiering, 2008). However, according to the company website, there are also steps being taken to create the recommendation system based on circulation data for other subscribing libraries.

Figure 4.2. BibTip- Enhanced OPAC: University of Karlsruhe

There were no concrete claims mas they were by Huddersfield University Library regarding the contribution of the recommender system in increasing the rate of borrowing. But Mönnich & Spiering(2008) explain the service is well accepted by users citing a user survey made at Karlsruhe between the years 2005 and 2006. According to this survey, most of the users found the recommender services to be very useful with average rate of 4.2 (on a scale of 1 (very poor) to 5 (very good)). In addition to that, the effort behind BibTip was seen as enriching library catalogs, giving OPACs a look of web 2.0 which is more appealing to the library's young generation of users who are well familiar of such features on web shops, YouTube and other Web 2.0 services. Moreover, it is also assumed that BibTip would be helpful in developing library holdings for it would show materials which are heavily used and materials which are least / never been used.

One notable problem of BibTip was the cold start problem (Gottwald and Koch, 2011). To find out more about this, Karlsruhe university library and the partner company which developed BibTip were contacted via e-mail. The company admitted that, as a self-learning and behavior-based system, BibTip requires learning period before local recommendations can be generated. The length of this learning period strongly depends on the relationship between the number of titles in the catalog and the search behavior of the users. They also added that they have had cases where they reached 10% coverage (in relation to the called titles of the last 30 days) within 2 months, but they have also had cases where it took 2 years to reach this mark.

However they claim that they have solved the cold start problem completely, because they are able to share the recommendations created within one library with other libraries. It is called recommendation interchange. This means, that they are able to fill-up the recommendation lists with recommendations created within other libraries (but only for locally available titles), but locally generated recommendations are always taking precedence when ranking the list of recommendations for a certain title. The more time passes by, the more locally generated recommendations will be populating the recommendations lists and only few foreign recommendations will be displayed among the recommendation lists.

4.4. bX Recommender of Ex Libris[11]

Ex Libris provides a recommender service called bX which is, according to the company website, *"based on data mining and structured analysis of usage data obtained from hundreds of research institutions worldwide".* It is developed in collaboration between ex Libris team and researchers Johan Bollen and Herbert Van Sompel at the Los Alamos National Laboratory[12].

bX was first announced on January 22, 2009 as "an exciting new Web 2.0 recommender service" via a press release by Ex Libris. It is also described as the first service to provide recommendations that are generated from aggregated usage data which is collected from participating institutions that already use Ex Libris's SFX link resolver. The bX service can be integrated with other systems such as Elsevier ScienceDirect and Scopus via an application programming interface (API)[13].

The bX recommendations are similar to the commercial "customers who bought this also bought…" recommendations. At that time it was described that sixteen institutions have begun testing the new service. According to Ex Libris group, the current number of institutions using bX has reached to 1000. bX Hot Articles was released recently as part of bX usage based services. This hot articles service identifies ten articles that researches have selected most in each discipline in recent weeks and also ten popular articles over all. This is a free service also available for mobile applications as Android and iPhone apps.

The main advantage of bX is its huge database. Recommendation services can start within the day the service is implemented in a library. It is easy to use as users don't need training to start benefiting from the service. Moreover, it is also mentioned as up-to-date because it includes current scholarly materials. Helping libraries fulfill their traditional mission of helping users find relevant materials, subscribing libraries can contribute their own usage data as part

[11] All information not cited in this section is obtained from Ex Libris group website
[12] www.exlibrisgroup.com
[13] http://libraryautomation.com/ltg-displaytext.pl?RC=16229

of contribution for enhancing the quality of recommendations for the subscribing society. But the contribution is optional (Ex Libris, 2011).

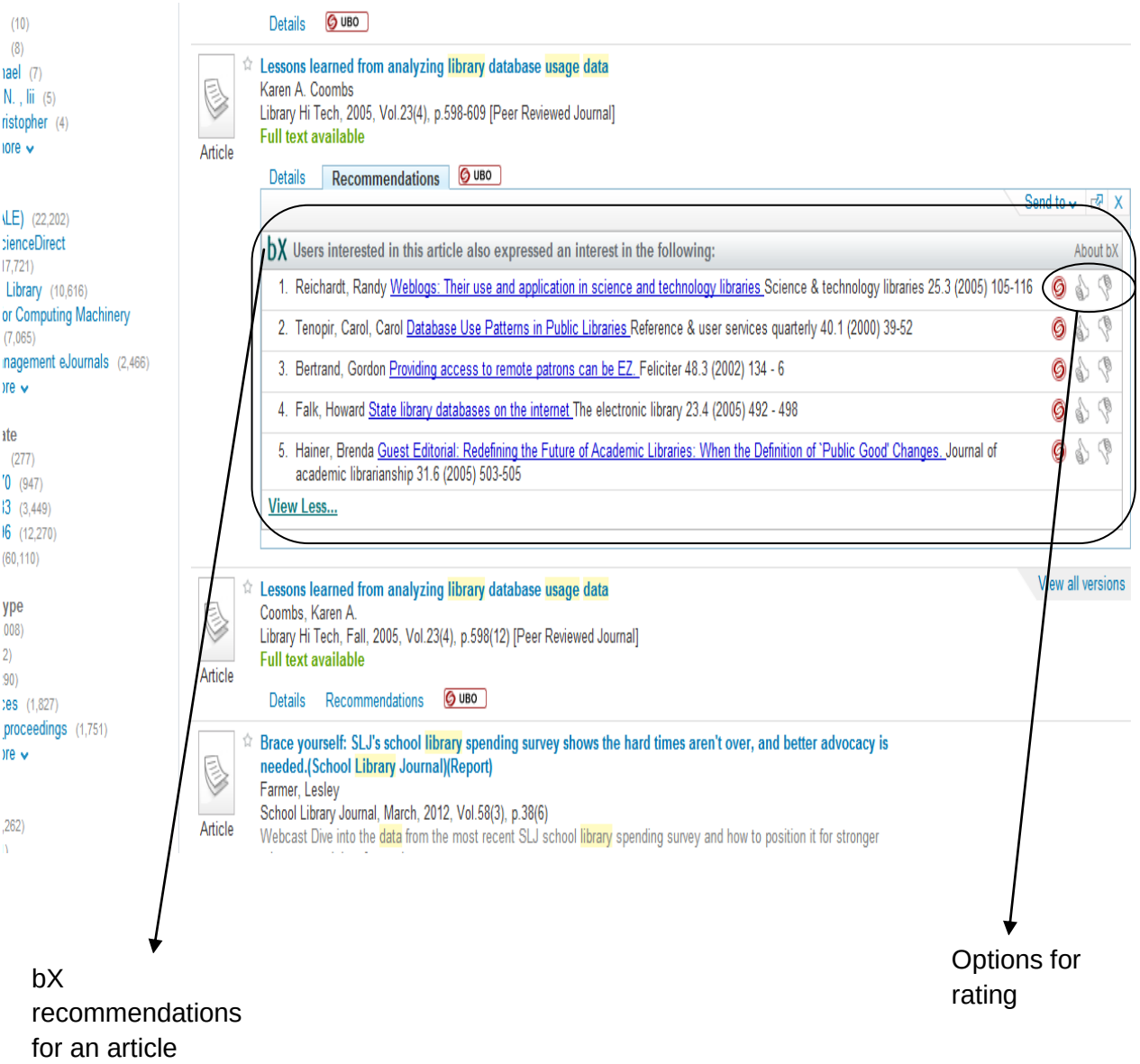

Figure 4.3. bX Recommendations

As a relatively new service, there seems shortage of literature on evaluating the actual use of bX by subscribing library users. But there was some tests made to know what users would think about the recommender service. Ponsford et al (2011) made a study on the use of Ex

Libris SFX menus including the bX recommender. The test was carried on 18 volunteer graduate students and two faculty members. One of the research questions was on whether users understand how bX recommendations work without further explanation and whether they find the recommendations relevant to their research needs. Then, the volunteers were asked to search for three articles and evaluate the relevance of bX recommendations. Most of them found out the recommendations were relevant and said they would follow them. Some of the respondents likened it to Google Scholar and Amazon. There were two respondents that have negative response to the relevancy of the search results and one of them, a teacher, voiced her concern that her students might rely on the first ten recommendations without committing more effort in finding best articles (Ponsford et al, 2011).

In another related study, Thomas et al (2011) interviewed eight undergraduate and two graduate students to get their view on the newly activated bX recommender system at the University of Waterloo's Primo Central. It was found out that the overall view of the students about the service was positive and they had also high expectations to find the recommended articles online. They found it easy, straightforward, and convenient but they were also able to find out some recommendations from a list which don't seem to fit in the group.

Apart the cases mentioned above, there seems lack of study on actual use of the recommendations based on a statistical data. There are indications of a growing momentum towards adoption of Ex-Libris by libraries (Ex Libris group, 2011) therefore it might be appropriate to conduct that kind of study to monitor the actual usage of the tool and get feedbacks and find ways of improving the service.

In order to get a glimpse of the use of bX recommendation usages, I contacted two university libraries in Norway namely University of Oslo and University of Bergen which are subscribing to the service as a part of the package of Ex Libris suite products. It was found out that in both cases the service was not promoted and its actual usage is very limited. However, It was possible to get some statistics that could help in showing the actual usage of the service in the libraries. According to the university of Bergen Library, the number of clicks for 2010 were found to be 5735 and for 2011, it was 6737. This means an average of 15.7 daily clicks in 2010 and 18. 5 in 2011. It was possible to get the usage statistics for 2010 at the University of Oslo

and it was presented in a more elaborated way, categorizing usage activities as 'requests' and 'clickthroughs' . *Requests* represent the number of times a user has clicked on a bX recommendation, and thus opened the SFX menu whereas *Clickthroughs* show the number of times the user has clicked on the SFX menu to open the recommended document. The total number of requests was 10158 (27.8 on average per day) and the total number of clickthroughs was 7792 (21.3 on average per day). Presenting the data as follows would illustrate what kind of usage data statistics to expect from the bX recommender system:

	Requests	**Click throughs**
***January**	*0	0
February	210	172
March	1621	1220
April	1736	1363
May	1407	1104
June	1073	854
July	540	426
August	848	663
***September**	5	0
October	1215	836
November	948	716
December	555	438
	10158	**7792**

Table 4.1. bX usage Statistics 2010: University of Oslo

* Data for January is missing and the figure for September is incorrect

The cases presented above present examples of incorporating recommender systems to library knowledge discovery tools and their characteristics can be summarized in the following table:

	Huddersfield University Library OPAC	PORE	BibTip	ExLibris bX
Recommendation type	Collaborative	Hybrid	Behavior based	Collaborative
Techniques	Data mining	ontology	Data mining	Data mining
Notable advantages	Increase circulation of books Provide alternatives for borrowing	Disambiguation of keywords. For example computer virus and biological viruses	Increase in borrowing Suitable for the young	Recommendation of scholarly journals, Presentation of "hot articles" in a given field of study, doesn't need training period
Notable problems	Limited only to popular and circulated books . subject to information decay	Needs periodic updating (e.g. once in a month) Subject to information decay	Cold start (as claimed, it is solved)	(though it needs further investigation) Irrelevant recommendations are spotted, old data problem
Tackling cold start problem	Data sharing via OpenURL protocol	"Default-Ontology" constructed beforehand for a new user.	Catalog crosslinking/ recommendation interchange	Harvest SFX
Contribution of the library	Developed the system (the library system division)	Partnership between the library and research units in the university	Installing and testing	Subscribe the service, user testing

Table 4.2. Summary of selected cases on recommender systems

4.5. MyLibrary

"*My Yahoo!, My CNN, My Bookmarks, MyThis and MyThat. Internet users have demanded a personal face to the World Wide Web, and Web portals and information providers have responded. Why not MyLibrary?"*(Cohen et al, 2000). This was one of the justifications given for the launch of MyLibrary applications at university libraries in the early 2000s.

Luce and Giacomo (2003) added that the changing nature of scientific research has added new expectations for libraries to support research. As they said, the future successful scientific library has to understand and aggressively respond to the needs of users engaged in the twenty first century research which is increasingly multidisciplinary, interdisciplinary, data intensive, and collaborative. Therefore they suggested that, as user needs become increasingly diverse, strategies to build digital libraries have to address the need of customization and personalization to accommodate the individual requirements and collaborative tools to work with the others. That was one of the visions when a MyLibrary application was launched at Los Alamos National Laboratory, one of the early adopters of MyLibrary functionality (Luce and Giacomo , 2003).

When users go to the physical library, the librarians would remember their users, how they have dealt with them in the past and that "stored data in their memory" would help them to satisfy their users' needs. What MyLibrary tried to do was to extend that personal touch to the digital realm (Storey, 2004). Its aim was to provide users to have their own personalized space where they can have customized view of their favorite resources, links, and tools for communication and collaboration (Cohen et al, 2000).Today, we see those features included on different digital library systems such as Dspace, CDs Invenio and Ex Libris products.

But recent trend shows that some universities that once started implementing MyLibrary personalization services have omitted it. For example, libraries at Virginia State University, University of Washington, California Polytechnic State Library, Cornell University, University of California and North Carolina University were among the early adopters of MyLibrary portals (Winter, 1999) but looking at their websites today, we can see that four of them no longer maintain the service. The following cases might help understand the reason behind.

The North Carolina State University Libraries was among those that stopped maintaining their MyLibrary portals. Writing about the five years history of MyLibrary functionality at the library portal, The head of The Natural Resources Library of the university libraries explains the idea behind implementing the MyLibrary@NCstate, as it was called , was to address the "information overload" experienced by library users. MyLibrary@NCstate, was both personalized and customizable. When users create new MyLibrary account, they were required to profile themselves to one of the disciplines in a controlled vocabulary roughly corresponding to the research and teaching fields in university. After they created the account, the software forwards them recommendations that include lists of resources in their disciplines, links to specialized sources and locations and also contacts of subject specialist librarians. It also allows them to add links to their favorite online or local resources. On the other hand, the customization feature requires users to interact with the software in adding links of any free resource on the web, adding resources from multiple disciplines and arranging them in a list, and even change the layout and colors of their pages(Ciccone, 2005).

The customization feature was later found to be cumbersome or time taking for the users, which might have discouraged them from using the system. The number of MyLibrary accounts , including non-active and rarely used accounts ,was found to be around 17 percent of the total population of users. The possible reasons for the decline of the usage of MyLibrary@SCstate were discussed as the time it takes to create and customize an account, fluidity of information ,i.e., the need to update the profiles as the students progress through their studies, lack of flexibility in the software, e.g., difficulty in turning features on and off as needed, creating, modifying, deleting categories, etc. and related software problems. Moreover, as the library website became more sophisticated with more functionalities than that of MyLibrary, the MyLibrary feature was being driven to irrelevance. For instance, as the library website enabled search across multiple databases and provided tools like E-Journal Finder, MyLibrary@NCstate couldn't provide such possibilities (Ciccone, 2005). As she explained, problems like those mentioned above have forced other libraries to adopt other functionalities such as course-related library pages that contain resources that meet the immediate needs of users related to current courses they are taking. She concluded that though the objectives of MyLibrary have been partially successful in meeting its objectives, its

failure was attributed to the limitation of the software. Therefore she noted that situation elicited investigation on ways of harnessing new technological developments to develop the next generation of MyLibrary, in order to realize its full potential(Ciccone, 2005).

In another case study made on the experience of personalization and customization, Shedlock et al (2010) presented the experience of personalization and customization at Galter Health sciences Library. They indicated that the library has been maintaining a web service called the Health SmartLibrary(HSL) From September 2003 to December 2008, with the goal of delivering health information, primarily full-text journal articles, directly and quickly to users, especially clinicians. The primary goal mentioned was saving users time in finding scholarly information to support quality health care. This service was modeled after commercial applications such as My Yahoo! and was built in response to users' needs to take control of their information environment. The HSL included tools such as *My E-Resources*, *Stay Current*, *Quick Search*, and *File Cabinet* to personalize and customize the users' experience at the library's website. My E-Resources brings most desired full text journals and other resources at the users' homepages according to their specialties they specified during time of registration or subscription to the service. This avoids the need of additional clicks to find the same materials through the library catalog or journal databases. Stay Current was current awareness service, Quick search was keyword search engine working across various library resources and File Cabinet was a unique folder designated for storage of links for a user's favorite resources (Shedlock et al, 2010).

The researchers said two web based surveys were conducted in 2005 and 2006. The objective of the first survey was to get comments on the overall satisfaction, perceptions, features, problems related to the service. The aim of the second survey was aimed at assessing the users knowledge of the tools included in the HSL. In addition to these surveys, the library was keeping track of usage of the tools by keeping statistics, i.e., records of customization activities of users. The overall result showed that users are receptive of the services but were largely reluctant to use the tools that require manual customization. The result also showed that more users adopted the automated applications (such as My E-Resources and Stay Current tools) than the others that require customization. The researchers recommended more research to determine why this is the case. But they acknowledged that pushing quality

information to the user, pushing the literature to the user-without the user spending time by searching, will have a great role in saving time of the user for other important engagements.

The researchers pointed out that this case presents a lesson as how librarians can help in the process. They said librarians can use the tools, their personal experience and selection skills to filter information available for users. Librarians can direct users' attention to quality information resources (Shedlock et al, 2010). The overall assessment of the two cases discussed above shows that more users tend to use the personalization features than the customization features of the MyLibrary applications.

Chapter Five

ANALYSIS

This chapter presents answers to the research questions as driven from the analysis of the presented cases and also as discussed under the literature review. Themes are extracted from the cases, supplemented with the literature review and finally presented under the research question they answer.

5.1. Research Question 1: Personalization Vs Library Functions

The first aim of this research was to see what functions of academic libraries can be enhanced by personalized services and how these services can be introduced in a way convenient to the user. The following paragraphs present the importance of such services to libraries as discovered in the research process and also presents choices for implementing them.

5.1.1. Value of personalization to Academic Libraries

The cases discussed indicate that personalization services can be used to enrich discovery tools like *library OPACs* (see 4.1 and 4.3). In the case of UH OPAC[14], it was mentioned that they took the inspiration from Amazon's "who bought this also bought…"feature. Comparison with Amazon and YouTube was also mentioned with BibTip to explain how the feature was accepted by the students who were already familiar with similar features on other web based services . Calling back the suggestions made by some authors to upgrade OPACs by adding web 2.0 features (see 3.7.2), we can see efforts like the above mentioned as enhancements of library OPACs in order to make them more valuable, presentable and attractive to users.

The issue of exposing items to user attention was mentioned in both cases mentioned above. For instance, the University of Huddersfield library claimed that circulation of books has increased after the addition of the recommender to their OPAC (though they also

[14] University of Huddersfield OPAC

acknowledge that further research would be required to support that claim). It is not explicitly mentioned whether the BibTip recommender has *boosted circulation* but the case they made for choosing behavioral data over circulation records reflects their intention of exposing more books, including the non-circulable reference materials, to the user attention.

We can also see an element of *reference service* being assisted by recommenders. For instance in the case of UH OPAC, it was mentioned that the recommenders might help a user to borrow alternative material if the one he wanted is not available for checkout. In the case of Ex Libris bX, we can see the expressed intention of the developers in designing the recommender to assist researchers in finding articles related to their research interests. If a recommender system can take up some roles of reference librarians that would mean librarians would save up time for other important activities. Time saving is also an advantage recommender systems present to the user (see 3.5). Personalization enables filtering and *pushing* relevant resources to the user so that the user doesn't need to spend time searching for those resources and save up the time instead for other important activities (see 4.5). However, it would also be necessary to take into consideration concerns by some saying such systems would discourage users from searching more important resources by themselves (see 3.5 ;4.4).

As it is indicated at the Copac Activity Data Project (where UH library is part of), librarians can use the suggestion lists generated by a recommender system for *collection development* purposes (Copac ,2012). Therefore we can add to this saying that, if personalization services help bringing the right materials into the focus of the right user, then academic librarians may need to consider them as extensions to their content strategies. This content strategy can be expressed as using personalized services to maximize usage of the acquired, processed and stored information resources.

As academic libraries repurposed their physical spaces to suit to the needs of their users, then as discussed under section 3.7.1, they can think about *repurposing their virtual spaces* to provide their users personalized areas where they can store their favorite resources that might include course blogs, internet links, library collections, etc. Such service can have special significance for virtual learners of virtual campuses (see 3.7.1). Users can use their

personalized spaces to create their own virtual communities of practice and share knowledge using variety of tools such as chat services and instant messaging. That was the objective of MyLibrary-like features (see 4.5). But these features suffered from little usage because of the reluctance of users to customize them, update them and log in time after time in order to use them (Nicholes and Mellinger, 2007; see 4.5) . This would entail the need of user study before designing library portals incorporating such features (see 3.6.3) and seek smarter ways of designing the portals to ensure the maximum possible user satisfaction. As the case of MyLibrary@NCstate suggests, if MyLibrary features are found to be inferior to the main university websites in terms of design, flexibility or functionality, then they will have a probability of being unused. This leads us to the next sub topic of this research question: how can personalized services be introduced in a way convenient to the user?

5.1.2. Adaptive or Adaptable?

Comparison of the two types of personalized services treated in this paper namely recommender services (UH OPAC, BibTip, PORE and Ex Libris bX) and personalized information spaces (MyLibrary) shows that the former are easier to use than the later. When we look at them closely, except for the probable need of logging in for authentication purposes, we can see that the recommenders need almost no effort from the user side. On the other hand, MyLibrary features need more effort and time and it looks that, unless users are convinced of the worth, they won't like spending much time on explicitly entering their preferences, updating their preferences, or customizing such features (see 4.5). As explained under section 1. 2, other researches too reflect the low usage of MyLibrary-like features. But looking deeper into the case shows that the automatic information *push* features incorporated in MyLibrary application would have a quicker chance of being adopted by users (see case of HSL under section 4.5).

Taking this back to implicit-explicit input discussions made under chapter 3, these cases show that the adaptive way or utilization of implicit user inputs seems to be the better option. The implicit inputs can be extracted from users' circulation records (UH OPAC, PORE) or from user interactions with the OPAC (BipTip, bX). As it can be seen from the figure 4.1., Document metadata can also extend the recommendations by pointing to materials written by the same

author or to the different editions of the same book (see 4.1). Comparison between implicit and explicit data also reveals explicit data are fluid or decay through time so they need to be updated periodically. That demands time and effort from the user side and that has attributed to the decline of MyLibrary like services (see 4.5) But if data harvested from user loan records for recommender purposes (see 4.1 and 4.2) can be considered as representing implicit user needs, then we can see an old data problem exists here too because such records age through time. These forces us to compare the behavioral data obtained through user's navigation (as used by BibTip) against the circulation record (as used by UH OPAC and PORE) to see which way is more suited to create personalized services that adapt themselves to the changing user needs (cf.3.6.2.3.) This presents a dilemma of whether to classify circulation records as sources of explicit or implicit user needs (more discussion on 5.4.2 and 5.4.3).

Though the cases discussed on this paper and the literature review show implicit input driven or adaptive systems could be better for designing tailored information services, studies mentioned under section 3.1. Show that the choice can be affected by other factors such as user cognitive style (field dependency or independency)and educational levels (see. 4.4.), among other things. This shows that it would be difficult to conclude which way is better and shows the need of user study before designing such services.

5.1.3. Recommenders: Collaborative or Content-based?

The other issue that could be addressed under this research question is what types of recommenders libraries might need to consider if they wish to add recommender functionalities on their OPACs. Though all of the recommenders we discussed employ the collaborative type of recommenders, researches show that collaborative and content-based recommenders have their own weaknesses and strengths (see 3.6.2). Therefore there are suggestions to mix them and create hybrid recommenders (see 3.6.2.3). Use of hybrid recommenders in also noted among the cases presented (see table 4.2). Moreover, the idea of Integrating social networking information with collaborative filtering is also forwarded as a solution for making recommendations more personalized (see 3.6.2.4).

5.2. Research Question 2: Usage Trend of Personalization Features

When we look at the overall library landscape based on the literature review, personalized services don't seem to top the list. Features such as MyLibrary and e-shelf which are seemingly abundant on many library management systems and scholarly journal websites are little used (see 1.2). They are even looked with suspicion by some librarians thinking that such tools could be outlets for compromising user privacy (Magi , 2010).

But as the usability tests made on Ex Libris bX recommender and SALT 2 Recommender (see 4.1, 4.3) show, users generally like the features. There were even cases where users likened the tools with Google and Amazon (see 4.1, 4.2, 4.3) which explains what users would like to see on their library and information systems. Going deeper into the group of users, there are cases which show personalized services are more preferred by higher level or postgraduate students than those at the lower levels (see 4.4). Those studies also showed that few respondents didn't like the features saying some recommendations are not relevant to the search result (see 4.4). A user survey made on BibTip also shows that users have rated the service highly. Therefore, libraries may consider promoting the services so more users may know about the tools and weigh in (see 4.4; McLaughlin, 2011). They could also take lesson from the case of MyLibrary on the need of designing flexible and user friendly applications.

5.3. Research Question 3: User Profiling and Privacy

5.3.1. User Profiling

As discussed under section 3.6, user profiles were being used to store explicit user preferences in the early days of selective dissemination of information. But as explained under section 3.4, libraries may not work anymore on collecting data that might relate to an individual. There is no discussion of user profiles in the cases analyzed on this paper except in the case of PORE and MyLibrary. The fact that circulation records are harvested at the Huddersfield University Library implies that there are profiles where user's loan history will be recorded and stored. However, as all records are anonymsed before being used for recommender purposes, there is no discussion of user profiling (see 4.1). In the case of BibTip,

the fact session ids are used instead of login ids demonstrates that user profiles are not a matter of discussion. Strict privacy requirements are explained in both cases as governing the design of the recommenders. Ex Libris library system allows users to have profiles but its bX recommender shares the same features like BibTip or UH OPAC in being collaborative recommender system that depends on 'collective wisdom' of crowds rather than attributes of individual users.

But in the case of PORE, user profiles have taken the form of personal ontologies, which are more elaborate representations of users. As discussed under literature review section 3.3, the use of ontology helps to solve vagueness that might result with traditional keyword based profiles. It helps to put a keyword into a context by defining its taxonomic and non-taxonomic relationships in a domain. The motive for PORE as explained in the case is "not to recommend a book on computer virus to a biologist". This demonstrates libraries can use their existing knowledge organization schemas as domain ontologies for creating personal ontologies for their users.

User profiles seem very important even though implementation of recommender services may not necessarily depend on them. The problem of most recommender systems is their failure to include *user context* which could be extracted from some personal information (see 3.6.3). User profiles provide that context and their integration with recommender systems makes a service more personalized (see 3.6.2.4). Ontologies can also be used to model the user context, broadening the user model to include all relationships in a domain (Stewart et al., 2004).

User profiles might be created with implicit and/or explicit data (see 3.3) but the discussions also show that those built with implicit data have the advantage of not aging over time so don't need updating (see 5.1.2). The explicit type has the advantage of being specific and precise in describing user interest but, as user needs change through time, such data needs periodic updating. Ontologies can also be a way for elaborate representation of user interests as shown with the PORE recommender system but they too are liable for information decay unless they are updated periodically (see 4.2).

In the cases that involve user profiles for personalization (MyLibrary and PORE), it was mentioned that users are given the opportunity of modifying their profiles. But as discussed under 5.1.2. This privilege comes with the demand of time and effort from the user side which might, by itself, be frustrating (see 5.1.2).

5.3.2. User Privacy

If users are asked to provide keywords for the materials they want to use, it would be realistic to assume that some of them may not respond or the others may list unrealistic preferences. But usage data, whether collected as circulation record or log data, is the real evidence libraries can have on their users changing needs so as to improve their services accordingly. Usage data can also give indication of which of the materials are most and least used therefore the data can be used as a tool for collection development. However, it would be worth noting that there are views suggesting past record can not necessarily tell what a user needs in the future (see 3.4).

Cases discussed under chapter 4 show us that usage data plays an important role in adding values to library tools such as OPACs and library portals, for instance, through addition of personalization and recommender services. The more information collected about the user, the more personalized the service will be (see 3.6.2.4). The need for privacy should also be addressed. But as discussed under the literature review and also presentation of selected cases, the need for privacy cannot be an obstacle for launching personalized services. Libraries can protect their users' privacies and at the same time provide personalized services by setting the appropriate policy frameworks and adopting technical capabilities for ensuring the security and privacy of users' data. The following methods, as discussed in the previous chapters, present options for balancing personalization and privacy:

- Let users know their activities are logged on and their data is solely used for personalizing services (Ferran, 2005) which means establishing thrust relationship with users.
- Adopt P3P standard of W3C (Neuhold, et al, 2003)

- Anonymising and aggregating data for use via data mining techniques (experience of Huddersfield University Library)
- Use session ids instead of login ids , usage of data mining techniques (BibTip)
- Implement client-server structure where personalized service runs on client side and retrieval service runs on the server side (Jing-Sen et.al ,2009)

5.4. Research Question 4: Usage Data Sources and Choices

5.4.1. Collecting Usage Data

Discussions made under sections 3.6.2.1 and 3.6.2.3 show that usage data serves as an evidence for linking a resource to a user. It is also an important base for recommendations.

The case of UH OPAC shows circulation records can be aggregated, anonymised and stored to be used by data mining applications for personalization purposes. Moreover it also shows that usage data on the electronic records can be collected from EZProxy logs, library link resolvers and library management systems. From the case of BipTip, we see users interaction with the OPAC can be captured and stored as behavioral pattern of a user and be used as a base to recommend items to him. From PORE, we can see that user's loan records can be used to build personal ontology to represent a user. Ex Libris' bX shows usage data from different institutions can be aggregated and stored at one place and be used to provide recommendation services to subscribing institutions. Usage data can also be acquired from publishers or their intermediaries though the level of statistics and the way of accessing them could be included as integral part of agreements while subscribing (Cole, 2000).

The cases analyzed and also the literature reviews show some problems with usage data that have to be dealt with. One of the problems is the old data problem which is revealed through a usability study made on SALT recommender system (similar with HO OPAC recommender) as discussed under section 4.1. If the circulation data is too old, the recommender might generate old or irrelevant materials. The case of PORE shows that the user data is updated once in a month to protect information decay. The other issue is the probability of getting faulty or inaccurate recommendation as reflected on a usability study on Ex Libris bX (see 4.1)

The other issue with the usage data is the need of critical mass of data. As reflected in the cases of UH OPAC, BibTip and Ex Libris bX, recommender systems need critical mass of data in order to work properly. Until the data reaches to that level, they experience the cold start problem. Libraries like the Huddersfield university library had the advantage of years' long efforts in accumulating usage data. But this might be challenging for others if they don't have the required mass of data. One very important lesson we can get from these cases is that libraries can join forces for central aggregation and sharing of usage and recommendation data via a certain protocol as it is being done under SALT 2 project of UK and BibTip (see 4.1; 4.3)

5.4.2. Circulation Data vs. Navigational Data

Having discussed methods of usage data collection in the manner presented above, the next question might be which type of usage data we should use. The circulation records as used by UH OPAC , PORE and to some degree by BibTip or browsing or navigational data as used by BibTip and bX?

Both groups of data have their own advantages and drawbacks. As it can be understood from the cases of BibTip and bX, activity data such as browsing and downloading history can be voluminous and have a better chance of portraying user needs and behavior. It would enable us to quickly get the volume of data required for data mining techniques to work on. But on the other hand, we need to be wary of its quality. As discussed under section 3.3., the term 'Usage' may need to be defined because browsing and clicking on an item may not correctly translate to usage. Section 3.3 also explains that user navigations can either be exploratory or goal oriented therefore it is difficult to say navigational data in totality would represent user's information behavior.

Circulation records on the other hand seem to reflect the exact need of a user at a given time but the problem is they take more time to reach to that critical mass. Moreover, processing them, for example, anonymising them (see case of UH OPAC), would be time taking. In addition to that, books with less copies may not have as much circulation history as those with multiple copies therefore the chance for them to appear among recommendations might be

low (see 4.3). Moreover as reflected in the case of PORE and also the activity data project in the UK, this type of data might get older through time and the recommendations it provides would become irrelevant. The other problem of circulation-record-based recommenders is that newly arrived materials and reference materials would not have chances of being included among the recommendations until they are well used.

5.4.3. Keywords, Concepts, Behavioral Data

The other approach for the choice of usage data and extending the discussion made above is to look through the types of records that would constitute a usage data. The following table shows possible types of usage data to be considered for use of recommender functionalities as it is learned from the cases discussed under chapter 4.

Usage data	**Advantage**	**Problems**	**Examples**
Keywords • Implicitly extracted from user's loan record • Explicitly entered by users	• Explicit in describing user needs	• Information decay (old data problem)	UH OPAC MyLibrary
Concepts • Keywords extracted from user's loan record then used to form personal ontology	• Solving synonymy and polysemy problems • Capture semantics of user profile	• Information decay: needs updating periodically	PORE
Behavioral data • Collected as users interact with library OPAC • Browsing history	• Attaining critical mass of data more quickly • No problem of information decay • Can also be information on the context of the user(Belkin and Zhang, 2004).	• Can we consider "clicks" as usage? • Distinguishing between exploratory and goal oriented navigation	BibTip Ex Libris bX

Table 5.1 Types of Usage Data with Examples

As it can be understood from the table above, combining the advantages of implicit data with concept based approach of ontology seems the ideal way of collecting and using usage data.

5.4. 6. Recall, Precision and Context

Recall and precision are catchwords which became important with the advent of information explosion and researches on information retrieval systems. Though the application of personalization in libraries nowadays is too limited, it would be worth asking whether we need to invoke those words. As it is mentioned under the discussion of Ex Libris' bX (section 4.4.), some of the recommendations may not be relevant to a search result they were meant to enrich. Moreover, it is worth calling back the polysemy and synonymy issues, where in the case of PORE, solution was given in the form of personal ontology. As it is found under the user testing undertaken by the SALT team (see section 4.1), some of the recommended items might be very old and out of date. Others might not be important for the user purposes (e.g. researches). This brings back the importance of context (as discussed under 3.2). These and other issues might be raised as the use of recommender systems increase. In this paper, there are two mechanisms identified which might be important to represent user context. The first one is use of personal ontology as used PORE and discussed under section 3.1 and the second one, as described by Belkin and Zhang (2004) is the use of implicit data which can also capture information on the user context (cf 5.3.1).

5.5. Research Question 5: Role of Academic Libraries

As we can see from the previous discussions, personalization systems can be built in-house in a library or in cooperation with computer science departments in a university (see the case of UH OPAC, PORE, MyLibrary), joint ventures between libraries and IT companies (see the case of BibTip) , or subscribed from external entity depending on the technical or financial capacity of the library (see the case of bX). Huddersfield University Library has their Library systems manager active on the design of their recommender system, and the feedback from National Chung Hsing University (see 4.2) shows that the library has a technical department that sees over the implementation and maintenance of PORE. Discussion on MyLibrary applications (section 4.3) show that university libraries can be involved in the design of personalized portals and follow-up the usage of those facilities. From the case of bX we can understand that librarians can have a role of subscribing for the services, installing the services and perform user tests to see what users like or don't like about the services, and promote the services if they find them useful (see also 3.6.1). In personalized systems that incorporate

push technologies (cf 1.2), Librarians can use their experience and tools available to push quality information to its right users (see 4.5). That will help in saving the user's time (cf 5.1.1).

As it can be learned from the years' long effort of aggregating and sharing circulation records by British university libraries (see 4.1), we can learn that a role of usage data curation can be added to the roles of librarians. Processing that data for use of personalized services would imply that librarians can take up the responsibility of user interest management.

The bX usability tests shown under section 4.4 and usability tests carried out by Copac (Copac 2012; section 4.1) show the need of evaluating recommenders and this can be another task for libraries. If current trends continue to hold, such services especially recommender services might become de facto on library OPACs. Therefore, the need of evaluation becomes obvious so as to improve and refine the services. Variety of techniques can be used to evaluate a personalized service including those mentioned on the Copac blog (see 4.1). For instance, focused group discussions with graduates and undergraduates can be made to know whether the recommendations have helped them in finding course materials. Interviews can be made with academics and teachers to see whether the recommendations have helped them to produce course reading lists. Librarians can also be interviewed if the recommenders support them in their tasks , such as collection development (copac, 2012).

Chapter Six

CONCLUSION

Following the changes taking place in academic institutions and the increasing trend towards individualized learning, it would be logical to say academic libraries, whose main purpose is supporting the goals and objectives of their parent organizations, should follow suit by offering individualized information services. As it is reflected in this thesis, technology has offered personalization and recommender tools for the realization of such services.

Following hints from literature about the importance of such services and the apparent little use of them by academic libraries, this research was prompted by the need of rediscovering the value of personalization for academic libraries and seeking answers for questions regarding its practical implementation. The research sought to discover functions of academic libraries personalization might improve , methods of introducing these services with the maximum possible user convenience and the role of librarians in running and maintaining the services. Moreover it sought to dig into components that make up any personalization tool, looking through the issues of usage data collection, user profiles maintenance and protection of user privacy. Selected cases were analyzed and themes extracted in the process were substantiated by extensive literature review to get answers to the research questions.

The study showed that academic libraries can employ personalization for two major purposes: enhancing their search systems such as OPACs by adding web 2.0 serendipity features and repurposing their virtual spaces. If OPACs are powered by recommender tools, there is a potential of exposing more resources to the user, boosting circulation of books and offering choices for users when materials they want are not available to borrow. Suggestion lists generated by the recommender tools can also serve as hints to librarians for decision making in collection development.

Large data sets are required for the recommender tools, e.g., data mining algorithms, to work which implies the need of collecting and storing usage data. While deciding on sources and types of usage data, the choice can be made between implicit and explicit user inputs or loan

records and behavioral data, weighing their advantages and disadvantages as discussed on this paper. Data mining techniques and ontology are among available tools for generating recommendations based on the stored data.

One issue that can be raised in connection to usage data is the critical mass of data needed for effective functioning of recommenders and means of overcoming the cold start problem. One solution presented to overcome this is central aggregation and sharing of usage data as experienced by BibTip and SALT 2 project in the UK. Those who choose to employ ontologies can follow the example of PORE in creating default ontologies based on the department and course information of students and refine the ontologies as the interaction between the users and the library progresses. The collaborative and hybrid (collaborative and content-based) recommenders are observed to be highly preferred in the cases discussed on this paper. Literature also supports the use of the hybrid types to improve performance of the recommenders. The other issue raised about usage data was old data problem which would result in faulty and irrelevant recommendations unless updated periodically.

Besides adding serendipity features to their OPACs, libraries may also consider offering personalized information spaces for their users. These services can be provided via library portals to give users their own virtual spaces where they can rearrange their favorite resources, share their favorite resources with others and have tools such as messaging or chatting services to share knowledge with others and on the way create virtual communities of practices. These tools would be important for all types of learners especially the virtual learners but researches also show that such services are suffering from declining usage. The effort and time they require for customization and periodic updating of user profiles was the main reasons mentioned.

There are underlying issues that govern or determine the success of such services. Besides the issues of usage data discussed earlier, the issue of user privacy has to be addressed. This study shows that the strong adherence of libraries to user privacy cannot be an obstacle as there are technical and non-technical procedures to choose from in order to keep the balance between privacy and personalization. Designing a policy framework for safeguarding privacy, adopting the W3C's P3P privacy protection standard are among the less technical procedures

while anonymising data , creating client –server tools where the personalization feature is administered by the user and information retrieval performed at the server side are among the technical methods discussed and proposed for safeguarding privacy while enjoying the benefits of personalization.

This study has found out that, among the selected cases for analysis, discussion about profiling academic library users is so little apparently due to privacy concerns. However, the discussion on harvesting loan records in cases such UH OPAC and PORE implies the utilization of user profiles .The literature review shows that it is difficult to enjoy the full benefits of personalization, which implies individualization, without user profiles. Elements a user profile has to include and the need of extending the item-user dimension of user profiles to add a third dimension called 'user context' are also discussed. Moreover, as a person might use different digital library systems, the concept of user profile interoperability to reconcile his/her profiles residing in different systems to achieve cross-digital library personalization is mentioned. In addition to that, the use of ontologies in transferring user profiles across systems has been discussed.

This study also highlights the task of librarians as user interest managers. On top of their traditional role as custodians of knowledge, personalization would enable them to be their user interest managers. This paper emphasized the need of managing usage data and demonstrated how librarians can go about harnessing it for mapping the right resource to the right user at the right time.

Though embrace of personalization by academic libraries is too little, there are evidences that most users would welcome such services. Studies discussed on this paper show users would prefer if the services are adaptive, taking little or no effort from them in customizing or adjusting the tools.

Chapter Seven

DISCUSSION

As it is mentioned at the start of this thesis, the value of personalization to digital libraries was expressed to the extent of justifying their existence. It was mentioned as a remedy for information explosion and a prerequisite for third generation digital libraries. There have been noble approaches and technical proposals forwarded for digital library personalization by different researchers at different times (Adomavicius and Tuzhilin, 2005). But the volume of researches made on their practical implementation in academic libraries is low (Neilson, 2008). Therefore, the aim of this study was to examine the practical situation of personalization in academic libraries and to rediscover its importance through analysis of five distinct cases representing different approaches to academic library personalization. Themes extracted in the process were enriched with conceptual literature review in order to get comprehensive answers to the research questions.

The study shows that personalization can be realized in academic libraries but the implementation issue seems complicated as it has to address underlying issues such as user data, usage data, and rules, regulations and practices governing the management of user and usage data. The purpose of this chapter is to elaborate on some important themes that stood out in the course of this study and attempt to point out directions for further research.

7.1. User Interest Management

Whether the personalization systems are produced in-house or purchased from an external entity, academic librarians can assume the role of making sure that such services are meaningful for their users. They can provide the logical platform, the context, the requirement analysis and specification part for the implementation of personalized services in their institutions. They would also assume the roles of conducting tests on the usability of the added services.

Librarians can also work on usage/usage data curation with the goal of improving and introducing user-aware services such as personalization and recommender services. In this study, usage data was found to be a very important element in library personalization and issues such as critical mass of data, information decay, and privacy were discovered along the discussion of usage data. This implies the need of data curation for addressing those issues and, as it can be derived from the experience of the selected cases, the process can go through the following stages:

- Usage data collection: identifying the sources to collect the data from, for instance, as discussed in chapter 4 under the first four cases, the sources can be circulation records, navigational history available from library management systems, link resolvers or EZproxy logs, usage data from publishers and database vendors, etc. Dealing with publishers and database vendors on the format and standard for accessing usage data of the electronic resources might be a challenging task.
- Usage data processing: it would include updating the data to prevent information decay (see PORE), anonymizing data to protect privacy (See UH OPAC, BibTip), etc. In case of implicit data, the task of identifying exploratory and goal oriented navigations might be challenging.
- Usage data storage: the decision would include how and where to store the data. This would include central aggregation and sharing of usage data as being done by Ex Libris and UK's SALT 2 project .
- Usage data usage and interchange: as shown in this study, the data can be used in designing recommender systems to enhance library search systems. The data can also be shared to libraries via a certain protocol to limit the cold start problem . experience can be drawn from recommendation sharing and interchange practices of UH OPAC and BibTip powered OPACs.

This evidently puts librarians between two interesting position. Safeguarding privacy of their users and looking to the inevitable future which requires utilization of user/usage data to realize more intelligent and user aware services. Past and ongoing researches show it is possible to do both. It is quite sensitive for librarians to be engaged with user data because of

the traditional sanctity of privacy in libraries. But the truth is, whether they like it or not, the future pulls them towards implementation of smart and intelligent libraries which cannot be realized without the consumption of user/usage data.

7.2. Privacy

Usage data is the soul of any meaningful personalized service as it provides the map between a resource and a user. But as indicated in the literature review (see 3.4), libraries tend to destroy evidence of usage after a material is returned with the motive of safeguarding user privacy. There are laws protecting them from keeping data related to transactions made between users and resources after the transactions have ended (see 3.4).

Libraries may have succeeded in protecting privacy when it comes to the usage of local materials. But what about the online resources they subscribe to? For instance, the Ex Libris bX recommender depends upon huge usage data collected as users from different subscribing institutions search for articles. Trends show that Ex Libris is being adopted by more and more libraries (see 4.4) which would have implications for future exploitation of usage data of users at different libraries. Publishers and their intermediaries produce usage statistics (which is also important to libraries) based on the data they collect from navigational activities of users at the subscribing institutions. The laws libraries use to defend privacy don't say so much about defending user's privacy on the online environment (see 3.4). Therefore, what steps should libraries take? Do they need to refine their privacy laws to avoid their user's navigation history from being utilized by online information services (Magi, 2012)? Do they need to delete old log files or seek ways of anonymizing them (Fifarek, 2002)?

User and usage data have high value in the market which can be demonstrated by the investments being poured on social networking sites (see 3.3). So why do libraries destroy such valuable treasure? Even though they are not for profit and cannot sell user data, can't they be creative enough in curating the data, while safeguarding privacy, and use them for creating intelligent services that ultimately benefit the users? (cf 5.3.2). It is doable as it is already being done in the cases discussed under sections 4.2 and 4.3. Therefore, what seems a viable option for libraries is to emulate the corporate world in understanding the value of

usage data and develop the capacity for managing and utilizing it to create intelligent information services. Libraries can design the service as a client-server application where usage data will be under the control of the client (cf. 5.3.2; 3.3), they can emulate social networks where users are responsible for managing their own data, they can anonymize the data, or they can invent other ways that can balance privacy and personalization. With ever increasing volume of digital content and information explosion, proliferation of digital libraries and propagation of needs for new generation libraries, personalization would be the way to go and usage data is an important part of it.

7.3. User Context

The other important theme that stood out in the course of this study was *user context*, which is an important concept not only in personalization but also in information retrieval research. Most personalization/ recommender items work in two dimensions, namely, item and user. But their problem is that they miss the third dimension which we call 'context' (Adomavicius & Tuzhilin, 2005) which, as discussed in this paper, is crucial for better understanding of user interests and better tailoring of information services.

Information services naturally require understanding of user needs which implies the utmost importance that should be given to a particular user's context. Maintaining user profiles would help to get this 'third dimension' but if librarians choose to refrain from maintaining user profiles, the other option could be utilization of social networks. Studies show that there are more than 900 million users of social networks such as Facebook, Linkdin, MySpace, Twitter, etc., that created a huge online repository of real identities (Wang et al, 2011; Toch et al, 2012). We can see now-a-days Facebook plugins incorporated on different websites to let users express whether the like the service or add if they have comments. We can also witness different Facebook apps that users willingly install in their accounts and share their information. Therefore, librarians may consider utilizing social networks to get themselves into the users context to know them better and better serve them (see 3.6.2.4). The issue of recall and precision is discussed under section 5.4.6. and addition of this third dimension (context) to personalization would help in solving problems associated with them.

7.4. Data Mining vs Ontology

Both bottom up, e.g., data mining (see 4.1, 4.3 and 4.4) and top down e.g., ontology (see 4.2) approaches can be used for implementing personalized services. The choice of explicit and implicit inputs is also on the table which requires weighing their particular advantages and disadvantages as discussed in the previous chapters. However, the main aim should be ensuring best user experience by designing the services that are user friendly to the possible best. As it is explained under section 3.6.3, it would be advisable if such efforts are preceded by a user study.

The data mining techniques function by clustering similar users and providing recommendations at the cluster level. This ensures privacy as there is no individualization involved but we cannot say such services are fully personalized. Individualization is achieved via user profiles and the question of whether it is possible to personalized information services to an individual level without having user profiles could be a subject for further research.

The ontological approach, on the other hand, provides a chance to create more elaborated user models . Therefore we can say services based on ontological user profiles are more personalized (see 3.2). But issues of user privacy and a possible problem of information decay require attention. The case of PORE has illustrated building personal ontologies would be possible in libraries where local collections are organized with the help of classification schemes like DDC or LC ,which can serve as domain ontologies where the personal ontologies can be mapped to. But can we extend this to include materials from external sources such as journal databases? This would require classifying online journals according to the classification scheme followed by the library and this also might require further research. under the literature review discussing user profiles and also in the case of PORE, one ontological approach for constructing user profiles is analysis of user's past borrowing history, extraction of concepts from the history and formation of ontologies representing user profiles. But such ontologies are prone to the problem of information decay (which is one of the main problems in explicit user profiles). In case of PORE it is mentioned that the user information is updated once in a month. This would suggest the need of extra effort in managing such user profiles.

Linking this to the discussion of implicit or behavior based user profiles (see 3.3; 5.1.2; 5.4.3), can we conclude that usage of the non-aging behavioral data will be superior? Can we create dynamic, self-updating, adaptive ontologies based on user's navigational history? This would trigger further examination and experimentation of the methodologies.

7.2. Significance of the Research

Given all what have been discussed above, this research provides some perspectives on how the issue of personalization can be approached in academic libraries. for those libraries that aspire to start some kind of personalized service, this research has carried important answers to important questions that can be asked during the planning phase of such service.

The discussions could also inspire future researches on user data curation and management, user interest management, management of trust and creating acceptable standards for sharing and reuse of usage data, design of personalized portals that protect privacy and at the same time provide fully personalized services, design of context-aware information services, adaptive ontological user models, etc. All major themes raised in the course of the study can themselves be future research tracks.

The future needs intelligent information systems and one thing we can be sure of is that libraries cannot escape this future. Therefore this paper would inspire academic librarians to keep adding values to the existing services they provide to achieve the ultimate goal of maximizing user satisfaction. The efforts of libraries discussed in chapter four can be a motivation as well as a lesson for academic libraries to work on adding values to their services. As recommender services get deeper roots in libraries, then issues related to data scarcity , quality of recommendations and usability can be among the daily engagements of academic librarians.

References

- Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*. IEEE Educational Activities Department. Retrieved from http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1423975

- Aghabozorgi, S. R., & Wah, T. Y. (2009). Recommender systems: incremental clustering on web log data. Proceedings of the 2nd International Conference on Interaction Sciences Information Technology Culture and Human (p. 812--818). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1656073

- Almazro, D., Shahatah, G., Albdulkarim, L., Kherees, M., Martinez, R., & Nzoukou, W. (2010). A Survey Paper on Recommender Systems. (E. Szczerbicki & N. T. Nguyen, Eds.)Arxiv preprint arXiv, abs/1006.5(5), 129-151. Springer. Retrieved from http://arxiv.org/abs/1006.5278

- Amato, G., & Straccia, U. (1999). User Profile Modeling and Applications to Digital Libraries. (S. Abiteboul & A.-M. Vercoustre, Eds.)Research and Advanced Technology for Digital Libraries, 1696, 184-197. Springer. Retrieved from http://www.springerlink.com/index/Q6E12FFNJTKELDKH.pdf

- Anderson, Byron. 2006. Keeping Up: SDI to RSS. Behavioral & Social Sciences Librarian 24, no. 2 (January): 113-117.

- Avancini, H., Candela, L., & Straccia, U. (2007). Recommenders in a personalized, collaborative digital library environment. Journal of Intelligent Information Systems, 28(3), 253-283. Springer. Retrieved from http://www.springerlink.com/index/10.1007/s10844-006-0010-3

- Awad, N. F., & Krishnan, M. S. (2006). The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. (H. Hippner & K. D. Wilde, Eds.)MIS Quarterly, 30(1), 13-28. MIS Quarterly & The Society for Information Management. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=19754858&site=ehost-live

- Badke, W. (2012). Personalization and Information Literacy. Online, 36(1), 47-49.

- Bailey, D. (2008). Transforming library service through information commons : case studies for the digital age. Chicago: American Library Association.

- Belkin, N. J., Muresan, G., & Zhang, X. (2004). Using User ' s Context for IR Personalization. ACM SIGIR 2004 Workshop on Information Retrieval in Context. (29

July 2004), 23-25. Retrieved from http://comminfo.rutgers.edu/~muresan/Publications/sigirWsBelkin2004.pdf.

- Boeije, H. (2002). A purposeful approach to the constant comparative method in the analysis of qualitative interviews. Springer Healthcare Communications, 36(4), 391.

- Bollen, J., Nelson, M. L., Geisler, G., & Araujo, R. (2007). Usage derived recommendations for a video digital library. Journal of Network and Computer Applications, 30(3), 1059-1083. Academic Press Ltd. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S1084804506000038

- Bollen, J., & Van De Sompel, H. (2006). An Architecture for the Aggregation and Analysis of Scholarly Usage Data. JCDL 06 Proceedings of the 6th ACMIEEECS joint conference on Digital libraries, (15 April 2010), 298-307. ACM Press. Retrieved from http://arxiv.org/abs/cs/0605113

- Bologna Declaration (1999), The European Higher Education Area. Joint Declaration of the European Ministers of Education, convened in Bologna, 19 June .

- Bryant, J., Matthews, G., & Walton, G. (2009). Academic libraries and social and learning space: A case study of Loughborough University Library, UK. *JOURNAL OF LIBRARIANSHIP AND INFORMATION SCIENCE*, *41*(1), 7-18. doi:10.1177/0961000608099895

- Brusilovsky, P., Cassel, L. N., Delcambre, L. M. L., Fox, E. A., Furuta, R., Garcia, D. D., Shipman III, F. M., et al. (2010). Social navigation for educational digital libraries. (N. Manouselis, H. Drachsler, K. Verbert, & O. C. Santos, Eds.)*Procedia Computer Science*, *1*(2), 2889-2897. Elsevier. Retrieved from http://dx.doi.org/10.1016/S1877-0509(10)00330-3

- Calhoun, K. 2006. "The Changing Nature of the Catalog and its Integration with Other Discovery Tools." Accessed April 20, 2012 from. http://www.loc.gov/catdir/calhoun-report-final.pdf.

- Cecelia, L., & Yang, Y. (2005). Leveraging Enterprise Technology for the Library Portal at the National University of Singapore. In Z. Chen, H. Chen, Q. Miao, Y. Fu, E. Fox, & E. Lim (Eds.), Digital Libraries: International Collaboration and Cross-Fertilization, Lecture Notes in Computer Science (Vol. 3334, pp. 783–802). Springer Berlin / Heidelberg. Retrieved from http://www.springerlink.com.ezproxy.tlu.ee/content/awxrdgrkle37hyrc/abstract/

- Chiu, W (2001). Website Personalization. IBM Developer Networks. Retrieved from http://www.ibm.com/developerworks/websphere/library/techarticles/hipods/personalize.html

- Chua, A. Y. K., and D. H. Goh. 2010. A Study of Web 2.0 Applications in Library Websites. *Library and InformationScience Research* 32(3): 203–211.

- Ciccone, K. (2005). MyLibrary@NCState: a library portal after five years. Journal of Library Administration, 43(1/2), 19-35. Retrieved from http://search.ebscohost.com.libaccess.sjlibrary.org/login.aspx?direct=true&db=ehh&AN=18940038&loginpage=Login.asp&site=ehost-live

- Cohen, S., Fereira, J., Horne, A., Kibbee, B., Mistlebauer, H., & Smith, A. (2000). MyLibrary: personalized electronic services in the Cornell University Library.DLib Magazine, 6(4). Retrieved from http://www.dlib.org/dlib/april00/mistlebauer/04mistlebauer.html

- Cole, L. (2000). Usage Data - The Academic Library Perspective. Serials: The Journal for the Serials Community, 13(2), 97–102.

- Coleho, H (2011). Web 2.0 in Academic Libraries in Portuguese Public Universities: A Longitudinal Study. Libri (61)4, 249-257.

- Copac (2011). User Feedback Results-Super 8. Retrieved from http://copac.ac.uk/innovations/activity-data/?p=216

- Copac (2012). Announcing the Copac Activity Data Project (otherwise known as SALT 2). Retrieved from http://copac.ac.uk/innovations/activity-data/?p=453

- Corbin, J., & Strauss, A. (2008). Basics of qualitative research (3rd Ed.). Los Angeles, CA

- Degemmis, M., Lops, P., & Semeraro, G. (2007). A content-collaborative recommender that exploits WordNet-based user profiles for neighborhood formation. *User Modeling and UserAdapted Interaction*, *17*(3), 217-255. Kluwer Academic Publishers. Retrieved from http://www.springerlink.com/index/10.1007/s11257-006-9023-4

- Dey, I. (1993). Qualitative data analysis: A user-friendly guide for social scientists. Nurse Researcher (Vol. 18, p. 285). Routledge. Retrieved from http://books.google.com.co/books?id=9acOAAAAQAAJ

- Estabrook, Alexia D.,& David L. Rothma. 2007. Applications of RSS - in Health Sciences Libraries. Medical Reference Services Quarterly 26, no. 1: 51. doi:10.1300/J115v26S01_04.

- Ex Libris Group (2011). bX Recommender Service: Usage Based Scholarly Recommendations. Retrieved from http://www.exlibrisgroup.com/category/bXRecommender#{C101F6E8-3FEF-4D60-BC13-024F56112BF9}

- Fan, W., Gordon, M. D., & Pathak, P. (2005). Effective profiling of consumer information retrieval needs: a unified framework and empirical comparison. Decision Support Systems, 40(2), 213-233. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0167923604000144

- Feagin, J., Orum, A., & Sjoberg, G. (Eds.). (1991). A case for case study. Chapel Hill, NC: University of North Carolina Press.

- Ferran, N., Casadesús, J., Krakowska, M., & Minguillón, J. (2007). Enriching e-learning metadata through digital library usage analysis. The Electronic Library, 25(2), 148-165. Retrieved from http://www.emeraldinsight.com/10.1108/02640470710741296

- Ferran, N., Mor, E., & Minguillón, J. (2005). Towards personalization in digital libraries through ontologies. Library Management, 26(4/5), 206-217. Emerald Group Publishing Limited. Retrieved from http://www.emeraldinsight.com/10.1108/01435120510596062

- Fifarek, A. (2002). Technology and privacy in the academic library. *Online Information Review*, *26*(6), 366-374. Retrieved from http://ejournals.ebsco.com.ezproxy.hio.no/Article.asp?ContributionID=4038352

- Foltz, P.W. & Dumais, S.T. (1992). Personalized information delivery: an analysis of information filtering methods. Communications of the ACM, v. 35, pp. 51-60.

- Frias-Martinez, E., Chen, S. Y., & Liu, X. (2009). Evaluation of a personalized digital library based on cognitive styles: adaptivity vs. adaptability. International Journal of Information Management, 29(1), 48-56. Elsevier. Retrieved from http://hdl.handle.net/2438/3195

- Frias-Martinez, E., Magoulas, G., Chen, S. Y., & Macredie, R. D. (2006). Automated user modeling for personalized digital libraries. International Journal of Information Management, 26(3), 234-248. Elsevier. Retrieved from http://www.sciencedirect.com/science/journal/02684012

- Gauch, S., Speretta, M., Chandramouli, A., & Micarelli, A. (2007). User Profiles for Personalized Information Access. (P. Brusilovsky, A. Kobsa, & W. Nejdl, Eds.)*Artificial Intelligence*, *4321*(3), 54 - 89. Springer. Retrieved from http://www.springerlink.com/index/y4g84202705577p3.pdf

- Geisler, G., McArthur, D., & Giersch, S. (2001). Developing recommendation services for a digital library with uncertain and changing data. Proceedings of the 1st ACMIEEECS joint conference on Digital libraries (pp. 199-200). ACM Press. Retrieved from http://portal.acm.org/citation.cfm?id=379483

- Gottwald, S., & Koch, T. (2011). Recommender Systems for Libraries [Survey and Forecast]. Machine Learning, 1-5. ACM.

- Gou, M and Zhao, Y (2008). An Architecture for Digital Library Information Service in An Ambient Intelligence Environment. Industrial Electronics and Applications 2008 ICIEA 2008 3rd IEEE Conference on (pp. 444-447).

- Guy, I., Zwerdling, N., Carmel, D., Ronen, I., Uziel, E., Yogev, S., & Ofek-Koifman, S. (2009). Personalized recommendation of social software items based on social relations. Proceedings of the third ACM conference on Recommender systems RecSys 09, New York,(7), 53. ACM Press. Retrieved from http://portal.acm.org/citation.cfm?doid=1639714.1639725

- Hanani, U. R. I., Shapira, B., & Shoval, P. (2001). Information filtering: Overview of issues, research and systems. User Modeling and UserAdapted Interaction,11(3), 203-259. Springer. Retrieved from http://www.springerlink.com/index/V25874116R1627P6.pdf

- Helfer, D. S. (2002) . Academic Entrepreneurship: The HBS Library Takes a Lesson from the School it Serves. Searcher 10 (4): 70–71.

- Hewitt-Taylor J (2001) Use of constant comparative analysis in qualitative research. Nursing Standard. 15, 42, 39-42. Date of acceptance: March 19 2001.

- Hensley, C.B. (1963). Selective Dissemination of Information (SDI): State of the Art in May, 1963. AFIPS '63 (Spring) Proceedings of the May 21-23, 1963, spring joint computer conference, 257-262. Retrieved from http://dl.acm.org/citation.cfm?id=1461584

- Im, I., & Harris, A. (2007). Does a one-size recommendation system fit all? the effectiveness of collaborative filtering based recommendation systems across different domains and search modes. ACM Transactions on Information Systems, 26(1), 4-es. ACM. Retrieved from http://portal.acm.org/citation.cfm?doid=1292591.1292595

- Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2011). Recommender Systems: An Introduction. Cambridge University Press.

- Jeevan, V. K. J. (2008). Library Personalization Systems: an Indian experience.IFLA Journal, 34(1), 72-83. Retrieved from http://eprints.rclis.org/15898/

- Jeevan, V. K. J., & Padhi, P. (2006). A selective review of research in content personalization. Library Review, 55(9), 556-586. Emerald Group Publishing Limited. Retrieved from http://www.emeraldinsight.com/10.1108/00242530610706761

- Jiang, X., & Tan, A.-H. (2009). Learning and inferencing in user ontology for personalized Semantic Web search. Information Sciences, 179(16), 2794-2808. Elsevier Inc. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0020025509001595

- Jin, Y. (2004). The development of the China Networked Digital Library of Theses and Dissertations. Online Information Review, 28(5), 367-370. Retrieved from http://www.emeraldinsight.com/journals.htm?issn=1468-4527&volume=28&issue=5&articleid=1506470&show=html

- Jing-sen, L., Guan-zhong, D., & Yu, L. (2008). A personalized retrieval system with preserving privacy. Industrial Electronics and Applications 2008 ICIEA 2008 3rd IEEE Conference on (pp. 2444-2448).

- Kaur, K. (2009) Marketing the Academic Library on the Web. *Library Management* 30(6/7): 454–468

- King, D. (2009). What is the Next Trend in Usage Statistics in Libraries? Journal of Electronic Resources Librarianship, 21(1), 4-14. Routledge. Retrieved from http://www.informaworld.com/openurl?genre=article&doi=10.1080/19411260902858276&magic=crossref

- Kohn, L. T (1997). Methods in Case Study Analysis. Technical paper. Retrieved from http://www.hschange.com/CONTENT/158/158.pdf

- Kolb, S. M. (2012). Grounded Theory and the Constant Comparative Method : Valid Research Strategies for Educators. Policy Studies, 3(1), 83-86.

- Konstan, J.A., Riedl, J.: Recommender systems: from algorithms to user experience. User Model. User-Adap. Inter. 22(1–2), 101–123 (2012)

- Lan-Xia, Fu (2010). Thoughts of the Construction Personalized Information Service System of University Libraries " Management and Service Science (MASS), 2010 International Conference on , vol., no., pp.1-4, 24-26 (2010) Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5576690&isnumber=5575325

- Lee, Y.-hsien, Cheng, T.-hsiang, & Hu, P. J.-hwa. (2009). Overcoming Small-size Training Set Problem in Content-based Recommendation : A Collaboration-based Training Set Expansion Approach. Design, 99-106.

- Liang, T., Yang, Y., Chen, D., & Ku, Y. (2008). A semantic-expansion approach to personalized knowledge recommendation. Decision Support Systems,45(3), 401-412. Elsevier. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0167923607000814

- Li Qian, Lizhen Liu, Yue Hu (2007) Analysis of Personalized Information Service System for Digital Libraries

- Liao, I.-E., Hsu, W.-C., Cheng, M.-S., & Chen, L.-P. (2010). A library recommender system based on a personal ontology model and collaborative filtering technique for English collections. The Electronic Library, 28(3), 386-400. Retrieved from http://www.emeraldinsight.com/10.1108/02640471011051972

- Liao, S.-C., Kao, K.-F., Liao, I.-E., Chen, H.-L., & Huang, S.-O. (2009). PORE: a personal ontology recommender system for digital libraries. The Electronic Library, 27(3), 496-508. Retrieved from http://www.emeraldinsight.com/10.1108/02640470910966925

- Liu, F., & Lee, H. J. (2010). Use of social network information to enhance collaborative filtering performance. Expert Systems with Applications, 37(7), 4772-4778. PERGAMON-ELSEVIER SCIENCE LTD. Retrieved from http://www.sciencedirect.com/science/article/B6V03-4XX23MG-P/2/22ff584b6ac134030d8b7f2d047aad91

- Liu, S. (2008). Engaging Users: The Future of Academic Library Web Sites. (H. Seiffert & G. Radnitzky, Eds.)College Research Libraries, 69(1), 6-27. American Library Association. Retrieved from http://crl.acrl.org/content/69/1/6

- Luce, R., & Di Giacomo, M. (2003). Personalized and collaborative digital library capabilities: responding to the changing nature of scientific research. Science technology libraries, 24(1-2), 135-152. Retrieved from http://cat.inist.fr/?aModele=afficheN&cpsidt=16082672

- Lynch, C. (2001). Personalization and Recommender Systems in the Larger Context : New Directions and Research Questions Recommender Systems and Information Discovery Personalization in a Distributed Information Environment. Second DELOS Network of Excellence Workshop on Personalisation and Recommender Systems in Digital Libraries, 84-88. Retrieved from http://www.ercim.eu/publication/ws-proceedings/DelNoe02/CliffordLynchAbstract.pdf

- MacManus, R (2009, January). 5 Problems of Recommender Systems. Retrieved from http://www.readwriteweb.com/archives/5_problems_of_recommender_systems.php

- MacManus, R (2009, January). Guide to Recommender systems. Retrieved from h http://www.readwriteweb.com/archives/recommender_systems.php

- Magi, T. J. (2010). A Content Analysis of Library Vendor Privacy Policies: Do They Meet Our Standards? College Research Libraries, 71(3), 1-41. Retrieved from http://crl.acrl.org/content/71/3/254.full.pdf+html

- McLaughlin, J. E. (2011). Personalization in Library Databases: Not Persuasive Enough? Library Hi Tech, 29(4), 4. Emerald Group Publishing Limited. Retrieved from http://www.emeraldinsight.com/10.1108/07378831111189723

- Mi, J., & Weng, C. (2008). Revitalizing t h e Library OPAC: Interface, Searching, and Display Challenges Jia Mi and Cathy weng. Information Technology and Libraries, 27(1), 5-22.

- Middleton, S. E., Alani, H., Shadbolt, N. R., & Roure, D. C. D. (2002). Exploiting Synergy Between Ontologies and Recommender Systems. System, 55, 10. Sementic Web Workshop 2002, WWW2002. Retrieved from http://oro.open.ac.uk/20058/

- Mobasher, B., Dai, H., Luo, T., Sun, Y., & Zhu, J. (2000). Integrating web usage and content mining for more effective personalization. (K. Bauknecht, S. K. Madria, & G. Pernul, Eds.)Electronic commerce and web, 1875, 165–176. Springer. Retrieved from http://www.springerlink.com/index/71HDXJKWMAVPBPDA.pdf

- Morita, M., & Shinoda, Y. (1994). Information filtering based on user behavior analysis and best match text retrieval. Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval, 272-281. Springer-Verlag New York, Inc. Retrieved from http://portal.acm.org/citation.cfm?id=188583

- Mönnich, M., & Spiering, M. (2008). Adding Value to the Library Catalog by Implementing a Recommendation System. DLib Magazine, 14(5/6), //wwwdliborg/dlib/may08/monnich/05monnichhtml. Retrieved from http://www.dlib.org/dlib/may08/monnich/05monnich.html

- Neilson, C.(2008). Current Awareness on a Shoe String -- RSS - at the HQC - PB - Routledge. Internet Reference Services Quarterly 13, no. 1: 57.doi:10.1300/J136v13n01_03.

- Neuhold, E., Niederée, C., & Stewart, A. (2003). Personalization in Digital Libraries – An Extended View. Proceedings of ICADL 2003 6th International Conference on Asian Digital Libraries, 1-16. Retrieved from http://www.springerlink.com/content/wc7kybhvtpuaf1y3

- Nichols, J., & Mellinger, M. (2007). Portals for Undergraduate Subject Searching: Are They Worth It? portal Libraries and the Academy, 7(4), 481-490. The Johns Hopkins University Press. Retrieved from http://muse.jhu.edu/content/crossref/journals/portal_libraries_and_the_academy/v007/7.4nichols.html

- Nika A. et al.,(2011). A Survey of Context Aware Cross Digital Library Personalization. Adaptive Multimedia Retrieval. Context, Exploration, and Fusion,6817(2011), 16-30. Springer. Retrieved from http://www.springerlink.com/content/a2p688150p718129/

- Park, D. H., Kim, H. K., & Kim, J. K. (2011). A Review and Classification of Recommender Systems Research. Social Science, 5, 290-294.

- Pera, M. S., & Ng, Y.-K. (2011). With a Little Help from My Friends: Generating Personalized Book Recommendations Using Data Extracted from a Social Website. Proceedings of the 2011 IEEEWICACM Joint Conference on Web Intelligent WI11.

- Perugini, S. and Gonçalves M. A. (2002). Recommendation and personalization: a survey. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.106.7207&rep=rep1&type=pdf.

- Perugini , & Gonçalves, A.(2002). Recommendation and Personalization: A Survey. Technical report cs.ir/0205059, Department of Computer Science, Virginia Tech, Available at http://xxx.lanl.gov/abs/cs.IR/0205059, 2002.Ponsford, B. C.,

- Ponsford, B. C., Stephens, J., & Sewell, R. R. (2011). Improving OpenURL Menus: User Testing of Revisions to SFX® Menus. Serials Review, 37(3), 162-170. Elsevier Inc. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0098791311000542

- Qian, L., Liu, L., & Hu, Y. (2007). Analysis of Personalized Information Service System for Digital Libraries. Information Technologies and Applications in Education, 2007. ISITAE '07. First IEEE International Symposium on (pp. 233 –237). doi:10.1109/ISITAE.2007.4409277

- Reisinger, D (2009).Top 10 movie recommendation engines.CNET. Retrieved from http://news.cnet.com/8301-17939_109-10200031-2.html

- Renda, E. M., & Straccia, U. (2005). A personalized collaborative Digital Library environment: a model and an application. Information Processing & Management, 41(1), 5-21. Pergamon Press, Inc. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0306457304000287

- Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to Recommender Systems Handbook. (F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor, Eds.)Recommender Systems Handbook (Vol. 40, pp. 1-35). Springer US. Retrieved from http://www.springerlink.com/index/10.1007/978-0-387-85820-3

- Ritz, T (n.d.). Modeling production of personalized information services and their delivery on multiple distribution channels. Retrieved from http://users.dimi.uniud.it/~stefano.mizzaro/AH2002/proceedings/pdfs/2ritz.pdf

- Samper, J., Castillo, P., Araujo, L., Merelo, J., Cordon, O., & Tricas, F. (2008). NectaRSS, an intelligent RSS feed reader. Journal of Network and Computer Applications, 31(4), 793-806. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S1084804507000495

- Savin-Baden, M. (2007). Learning spaces: Creating opportunities for knowledge creation in academic life. New York: McGraw-Hill.

- Schafer, J. B., Konstan, J., & Riedi, J. (1999). Recommender systems in e-commerce. Proceedings of the 1st ACM conference on Electronic commerce EC 99, 110(21), 158-166. ACM Press. Retrieved from http://portal.acm.org/citation.cfm?doid=336992.337035

- Shedlock, J., Frisque, M., Hunt, S., Walton, L., Handler, J., & Gillam, M. (2010). Case study:. Journal of the Medical Library Association JMLA, 98(2), 98-104. Medical Library Association. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2859266/

- Smeaton, A. F., & Callan, J. (2005). Personalisation and recommender systems in digital libraries. International Journal on Digital Libraries, 5(4), 299-308. NSF-EU. Retrieved from http://www.springerlink.com/index/10.1007/s00799-004-0100-1

- Smith, Geri; Street, Kelvin and Wales, Tim (2007). An online library service for Open University MBA Alumni: challenges and opportunities. Journal of Librarianship and Information Science, 39(3), pp. 162–176.

- Stephens, J., & Sewell, R. R. (2011). Improving OpenURL Menus: User Testing of Revisions to SFX® Menus. Serials Review, 37(3), 162-170. Elsevier Inc. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0098791311000542

- Stewart, A., Niederée, C., Mehta, B., Hemmje, M., & Neuhold, E. (2004). Extending Your Neighborhood-Relationship-Based Recommendations Using Your Personal Web Context. Digital Libraries: International Collaboration and Cross-Fertilization (Vol. 3334, pp. 523–532). Retrieved from http://140.234.0.9:8080/EPSessionID=178ae7dc32f49c4becc52af4279d01f/EPHost=www.springerlink.com/EPPath/content/7jam5n6qj07pvx5j/

- Storey, T(2004). Personalizing The Web: Libraries Look to Balance, Technology, Cost and Useulness. OCLC Newsletter. No 266. Retrieved from http://www.oclc.org/news/publications/newsletters/oclc/2004/266/personalize.html

- Su, X., & Khoshgoftaar, T. M. (2009). A Survey of Collaborative Filtering Techniques. Advances in Artificial Intelligence, 2009(Section 3), 1-20. Hindawi Publishing Corp. Retrieved from http://www.hindawi.com/journals/aai/2009/421425.html

- Tedd, L. A., & Large, A. (2005) Digital Libraries Principles and Practices in a Global Environment. Munchen: K. G. Saur

- Tellis, W (1997). Application of a Case Study Methodology. The Qualitative Report 3(3) retrieved from http://www.nova.edu/ssss/QR/QR3-3/tellis2.html

- Thomas,D., Greenberg, A.,& Calarco, P.(2011, September). Scholarly Usage Based Recommendations. PowerPoint Presentation. Retrieved from. http://igelu.org/wp-content/uploads/2011/09/bx_igelu_presentation_updated_september-13.pdf

- Toch, E., Wang, Y., & Cranor, L. F. (2012). Personalization and privacy: a survey of privacy risks and remedies in personalization-based systems. User Modeling and UserAdapted Interaction, 22(1-2), 203-220. Retrieved from http://www.springerlink.com/index/10.1007/s11257-011-9110-z

- Tv Genius (2011). An Integrated Approach to TV & VOD Recommendations. Retrieved from http://www.itvt.com/blog/how-overcome-barriers-personalised-tv

- UNESCO (2003). Digital Libraries in Education: Analytical Survey. Moscow: UNESCO Institute for Information Technology for Education.

- VanUllen, M. K., & Germain, C. A. (2002). Business as usual: Amazon.com and the academic library. The Journal of Academic Librarianship, Volume 28(52002), 319-324.

- Vellino, A. (2010). A comparison between usage-based and citation-based methods for recommending scholarly research articles. Proceedings of the American Society for Information Science and Technology, 47(1), 110. Wiley Online Library. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/meet.14504701330/full

- Wakeling, S., Clough, P., Sen, B., & Connaway, L. S. (2012). Readers who borrowed this also borrowed?…?": recommender systems in UK libraries. Library Hi Tech, 30(1), 134-150. Emerald Group Publishing Limited. Retrieved from http://www.emeraldinsight.com/10.1108/07378831211213265

- Wallis, K. & Kroski, E.(2009). The Next Generation of OPACs in Academic Libraries. [Accessed: 18-04-2011]. Available on Internet: http://eprints.rclis.org/bitstream/10760/13718/1/Term_paper_pdf.pdf

- Wang, Y., Norcie, G., & Cranor, L. F. (2011). Who Is Concerned about What? A Study of American, Chinese and Indian Users' Privacy Concerns on Social Network Sites. (J. M. McCune, B. Balacheff, A. Perrig, A.-R. Sadeghi, A. Sasse, & Y. Beres, Eds.)Trust and Trustworthy Computing. Springer Berlin Heidelberg. Retrieved from http://www.springerlink.com/index/LU75324R42234601.pdf

- Winter, K. (1999). "MyLibrary" can help your library. American Libraries, 30(7), 65-67.

- Wu, Y. D., & Liu, M. (2001). Content management and the future of academic libraries. The Electronic Library, 19(6), 432-440. Retrieved from http://www.emeraldinsight.com/10.1108/02640470110412044

- Yang, S. Q., & Hofmann, M. A. (2011). Next generation or current generation?: A study of the OPACs of 260 academic libraries in the USA and Canada. Library Hi Tech, 29(2), 266-300. Emerald Group Publishing Limited. Retrieved from http://www.emeraldinsight.com/10.1108/07378831111138170

- Jin, Y., & Peng, J. (2009). Information portal development and practice at Shanghai Jiao Tong University Library. Online Information Review, 33(3), 537-547. Retrieved from http://www.emeraldinsight.com/10.1108/14684520910969943