Leveraging student course enrollment data to infuse personalization in a library website

Ian Chan

Library Systems, California State University San Marcos, San Marcos,

California, USA

Abstract

Purpose – The purpose of this paper is to describe the benefits of integrating personalization within a library web site and presents methodology for achieving this goal within an academic setting.
Design/methodology/approach – The project documented in this study explores the use of student course enrollment data as the basis for creating a personalized library web site. Off-the-shelf, open source applications are used in conjunction with existing university data to deliver a final product that offers an enhanced user experience for the university community.
Findings – Adaptive personalization is increasingly commonplace on the web. Academic libraries have a unique source of existing data that offers the potential of adding personalization to the library web site. At present, the personalization of library online services remains largely unexplored. This project illustrates one relatively low-cost method to help libraries interested in creating personalized web sites.
Practical implications – This paper provides a guide for libraries interested in the implementation of personalization within their web sites.
Originality/value – The project described in this case study is highly unique within libraries. The paper outlines the feasibility and technical requirements associated with using course enrollment data to add personalized content to a library web site.
Keywords User interfaces, Personalization, Human computer interaction, User-centered design, Student-centered library, User experience design
Paper type Case study

Introduction

While academic libraries have previously offered customizable web sites, adaptive web site personalization through the integration of student data remains unexplored. This case study describes the implementation of a system that utilizes student course enrollment data to create adaptive personalization within the context of an academic library web site. The study presents the design, development, implementation, and assessment of such a system.

The challenge of building a user-friendly academic library web site

Academic libraries continue to develop their web sites as gateways to information resources, research help, and services. This investment has led to a great deal of research focusing on the assessment and improvement of academic library web sites (Brantley et al., 2006; Kim, 2011). A number of these studies show that users continue to find academic library web sites complex and difficult to use. In focus groups composed of students and faculty from two universities, many "participants commented that library websites are overly complex and hard to navigate, and that a simplified portal designed to meet individual needs would be welcome" (Munro and McLure, 2010).

Participants in a 2011 study of a university library web site "perceived the usability of the [library's] website design to be challenging" (Kim, 2011).

Recommendations aimed at improving library web sites have increasingly called for libraries to take a user-centric approach in designing their sites. A 2008 article reviewing 111 Association of Research Libraries web sites found "the universe of information presented on academic library homepages still focuses on library functions, requires numerous pathways for access, has overwhelming options, and takes a 'one-design-for-all' approach that fails to recognize users as individuals."

The author recommends libraries offer each user a "personal library space," based on that user's profile, to reduce information overload and present "library resources in a targeted and customized manner" (Liu, 2008). Somerville and Brar (2009) suggest that academic libraries should re-design their web sites and online services from a user-centric rather than library-centric perspective.

Studies of factors impacting the use of academic library web sites also point to a positive correlation between the perceived-ease-of-use (PEOU) of a site and the future intentions to use the site. Libraries that improve the PEOU of their sites are likely to increase the desire of students to use the library's web site (Heinrichs et al., 2007; Kim, 2010). A study of user perceptions toward university library web sites recommends practitioners "design user-focused library websites that enhance the usability of [the university library website] and provide customized services to different user groups in order to increase usage" (Kim, 2011).

Using personalization to create a user-centric library web site

Personalization can be defined as "the ways in which information and services can be tailored to match the unique and specific needs of an individual or a community. This is achieved by adapting presentation, content, and/or services based on a person's task, background, history, device, information needs, location, etc. essentially the user's context" (Smeaton and Callan, 2005, pp. 299-300). Frias-Martinez et al., 2009 describe two major approaches to personalization: adaptability and adaptivity.

Research on the potential benefits of personalization systems continues to grow (Park et al., 2012; Sunikka and Bragge, 2012). Improving web site ease-of-use and reducing information overload are two oft-cited benefits of personalization (Ketchell, 2000; Kumar and Benbasat, 2006; Liu, 2008; Wang and Yen, 2010; Lee and Cranage, 2011). Research conducted by Liang et al. (2007) indicates that "reducing information overload is the most important concern for users in seeking information and that personalized recommendations can perform well when users use the media to seek specific information." A study by Porter (2011) on undergraduate research strategies suggests that personalized recommendations would help students locate relevant databases.

In the early 2000s, a number of academic libraries implemented web site personalization systems. A review of web site personalization initiatives from that time period found most favored the use of adaptable or user-customized systems (Jeevan and Padhi, 2006). A number of academic libraries used MyLibrary, a web-based content management system (CMS) that gave users the ability to customize the layout and content of a library web site (Morgan, 1999; Cohen et al., 2000; Ghaphery, 2002; Gibbons, 2003).

However, research on how users perceived the adaptable personalization systems offered by libraries has shown those systems encountered limited success. Gibbons (2003) Student course enrollment data and Ciccone (2005) analyzed usage of the library web sites at their academic institutions and discovered that only a small percentage of users choose to use the customization functions. Responses to a survey conducted by librarians at Oregon State University, indicated that "supposedly tech-savvy students were unenthusiastic about a customizable [library] portal" and that "if the burden of customization [was] on them, only a small percentage would take the time to set up and use those features" (Nichols and Mellinger, 2007). In a 2011 study of academic library users, participants reported that it was difficult to customize the library web site and many did not utilize

the customization functions (Kim, 2011).

The limited acceptance of customizable library web sites may be due to their use of the adaptable approach to personalization. In contrast, the adaptive approach is generally perceived to be more favorable among users. Results of a five-year study at the Galter Health Sciences Library "indicated that users were receptive to personalized resource selection and that the automated application of specialty-based, personalized HSLs was more frequently adopted than manual customization by users" (Shedlock et al., 2010). This supports the results of earlier studies where users preferred adaptive over adaptable personalization (Frias-Martinez et al., 2006, 2009).

The adaptive or system-driven approach to personalization utilizes automation to produce user profiles that are based on the analysis of user interests and behavior (Frias-Martinez et al., 2006). In some cases, content relevance is dependent on data provided by the user as part of an account registration process. The personalization system correlates each user's data with relevant products, services, and information in order to generate recommendations. The adaptive approach might also incorporate analysis of a user's browsing and purchasing patterns (Forsati and Meybodi, 2010). Some adaptive systems include functions that make adjustments based on implicit user feedback (Pu et al., 2012).

Major challenges in the implementation of adaptive personalization include the development of an appropriate user model and accurate analysis of user browsing data. "There is considerable difficulty in getting real and correct user interests and mapping them effectively into the products and services offered by the library. Also, the interests of users keep on changing continuously" (Sirisha et al., 2009). Unless a user logs-in with the system or has previous browsing history on a site, the system will not have a basis from which to generate content relevance. Even if a user authenticates or allows tracking of their browsing history, their goals and interests may differ between or within visits.

Research on the use of adaptive personalization in the context of academic library web sites and digital libraries is heavily focused on deriving user interests through the analysis of user searches and click-throughs (Sunikka and Bragge, 2012). The integration of user data held within the student information systems (SIS) of academic institutions is a heretofore unexplored area within the literature.

Applying an adaptive personalization solution utilizing student course enrollment data

As with other academic libraries, a primary goal of the California State University San Marcos (CSUSM) Library is to design efficient pathways to information resources. To achieve this goal, the Library focusses on reducing barriers to resources and ensuring consistent, user-friendly information-finding tools. The Library's web site plays an essential role in this effort as it provides the primary means by which users locate and retrieve information. To support this mission, the Library invests significant resources toward creating a positive and fulfilling online research experience for its users.

The CSUSM Library also focusses heavily on providing course-integrated information literacy instruction and personalized research assistance. This instructional strategy places an emphasis on meeting the specific research needs of individual students. Library web site personalization was perceived as a strategy that might improve the learning experience as it would focus on each student's area of study. With the goal of providing a better, targeted research experience for students, the Library initiated a project to investigate the use of personalization on its web site.

As described in the preceding sections, research on the use of personalization in libraries clearly indicated that an adaptive approach was more likely to achieve success. However, the development of a data-mining system for automated personalization was beyond the scope of the Library's personnel resources.

In addition, an approach based purely on the data-mining of web site usage would have been difficult and unlikely to succeed for two reasons: students often enroll in multiple classes of different subjects and the development of an algorithm to accurately analyze the vast amount of usage data would have been very complex. This need to consider alternative methods of adaptive personalization led to the realization that academic libraries have access to a tremendous amount of information on their users: course enrollment data. Using this data as the basis for personalization would allow us to generate user profiles without requiring implicit input from our students. Content on the library web site could then be associated with each user's profile based on relevance to the student's courses. With this data model in mind, we choose to explore the use of course enrollment information as the foundation of our personalization strategy.

Project outcomes and requirements

The overall goal of the personalization project was to enhance the research experience and create a more user-centric library web site. The primary outcome was to develop a system that automated the process of connecting users with the library resources most relevant to their research needs. It would offer simpler pathways for accessing online resources and enrich the overall user experience.

In order to maintain data security and user privacy, it was critical that the system utilize a reliable and secure user account management system. To minimize maintenance overhead for the Library and for the CSUSM Instructional and Information Technology Support group, the personalization system would have to leverage existing user data stores. The system would need to visually highlight resources and research guides relevant to the researcher's subjects of interest. It would also display the appropriate librarian profiles within a user's subjects of interest. Users would not be required to create accounts, self-select subjects areas, or login to additional systems. Lastly, it was essential that the system utilize existing open source applications and that it would not require significant investment of programming resources.

Developing a data model

The effectiveness of a personalization system is dependent on the acquisition of pertinent user data and utilizing that information within the framework of a data model. The process can be summarized as follows: "(a) the collection of Web data, (b) the modeling and categorization of these data (preprocessing phase), (c) the analysis of the collected data, and (d) the determination of the actions that should be performed" (Eirinaki and Vazirgiannis, 2003, p. 4).

The initial phase of the project focused on building a reliable transfer of course enrollment data and developing methods of applying that data to personalization. For the project's data requirements, it was necessary to develop a recurring process that would make available the course numbers and instructor names associated with each student's course enrollments. An existing process provided an up-to-date extract of student account information that was transferred on a regular basis from the SIS to the integrated library system (ILS). By expanding this process to include student course enrollment data from the SIS, we were able to meet the data requirements of our personalization system.

To associate library resources with individual students, we needed to cross-reference the course enrollment of each student with a set of primary subject headings. All of the Library's resources are organized into 30 primary subjects and each subject corresponds to one of the disciplines associated with the courses offered by CSUSM. To create the cross-references between courses and library primary subject areas, we setup a database table to link all course numbers with their relevant library subject areas. This table matches the departmental prefix of each course number with the associated library subject. For example, all courses prefixed with "MRKT" are identifiable as relevant to the areas of marketing and business.

In the personalization system, each user profile would include the library subject areas most relevant to the student's course enrollments. By cross-referencing subjects in the user's profile with those associated with the resources, guides, and librarian profiles on the Library web site, the system would be able to generate recommendations.

Selecting Drupal as the CMS for supporting personalization

Building a homegrown system with user profiles, account management, resource metadata, and content management was beyond the scope of the Library's resources. It was more cost-effective and sustainable to customize and extend an existing web-based CMS that already incorporated the requisite functionality.

Drupal (http://drupal.org) is an open-source CMS offering a combination of features that make it an ideal candidate for building a web site personalization system. Those features include user accounts, highly extensible content types, and a robust taxonomy system. In addition, it offers a powerful, browser-based tool for creating and displaying relational database queries. Its extensive list of add-on modules meant we would not have to invest resources toward developing new functionality to meet our project requirements.

The following is a list of requirements meet by Drupal:

- ability to import external data and map to internal content fields;
- integration with external Lightweight Directory Access Protocol (LDAP) authentication systems;
- ability to create and manage a highly varied set of content types, each with its own set of data fields;
- ability to create and display complex queries without extensive programming knowledge;
- a flexible and secure user profile system; and
- a highly configurable presentation layer.

Based on these requirements, Drupal was selected as the CMS for the personalization system. When Drupal was implemented as the Library's web site CMS, all of the Library's e-resource profiles, research guides, and librarian profiles were created as "nodes" of content within the system. In Drupal, a "node" refers to a single entity of content that may contain any number of fields. Within the context of the node system, it is possible to designate any number of node types, each with its own configuration and set of fields. Each node type may have any number of fields. Creating and managing node types, node fields, and nodes is accomplished via Drupal's web-based administrative interface. Prior to version 7, managing the fields of a node type was enabled via an add-on module known as the Content Construction Kit (CCK). In Drupal 7, the node management features of the CCK module are incorporated within the core application.

Within the CSUSM Library web site, the e-resource node type is an example of how we use Drupal's node-based content system. This content type is used to store information describing the attributes of online resources such as databases, e-books, and e-journals. The fields in the e-resource nodes hold information such as database URL, authentication method, full-text availability, dates of coverage, and more. User profiles, research guides, course guides, and librarian profiles are also nodes of information within Drupal. Building relevance between content and users is possible because each node is associated with subject terms entered into the system's central taxonomy.

An essential element of a personalized web site is the privacy and security of user data. The Drupal user account system restricts public access to user data while allowing users to access and update their own profiles. For added security and ease of maintenance, we do not store student ID numbers,

residential addresses, passwords, and phone numbers in Drupal. While Drupal is capable of securing that type of information, those elements of user data provided no added benefit to the personalization system.

Importing user data

Developing an efficient and sustainable process for loading student enrollment data into Drupal was critical to the project. The data formatting and import requirements of Drupal made this the most challenging aspect of building the personalization system. The following section highlights several key elements of the data import process.

Prior to loading the user data exported from the SIS, that data must be formatted to match the table schema of Drupal's MySQL database. It is also necessary to add the appropriate library subject areas to each user's data within the exported SIS data. This is a multi-step process that occurs outside of the Drupal system and utilizes a combination of regex queries, MySQL queries, and PHP scripting. The regex queries ensure that the fields and columns within the exported table of SIS data match those in Drupal's user tables. After the regex queries normalize the SIS data for use in Drupal, the data is imported into a standalone MySQL database. The PHP script runs several MySQL queries to analyze each user's enrollments and associate each of the student's classes with the relevant library subject areas. The script also checks to see if a student's enrollment data has changed. This allows us to filter out unchanged records that do not require updating in Drupal. The data loaded to the MySQL database is reused for other aspects of this project and for a related course reserves system.

Importing the user enrollment data was complicated by the fact that Drupal 6 defines user accounts and content nodes as dissimilar data entity types. This was a barrier to the personalization process because it prevented direct association of user accounts with taxonomy terms. The content profile add-on module addresses this issue by attaching a single content node to each user account. By limiting each user account to a single attached content node, it becomes possible to associate subject terms with individual users. Within the personalization system, these user profile nodes store the course enrollment data for each CSUSM student.

However, this does create the need to use two different processes for importing student data into Drupal. One process imports basic user account data, such as name and e-mail, and automatically creates or updates user accounts. The second imports the subject data relevant to each student's course enrollments and creates or updates each student's profile. While the two processes require separate import methods, they both utilize the same data file. The Drupal modules used for the data import are user import and node import. In each of those modules, the import process includes uploading the data file, matching the file columns to fields, and automated batch-processing of data. E-mail address is used as the unique identifier to match imported data with existing accounts and profiles. If the user does not have an account or profile within the system, those are automatically created. The node import update module extends node import to ensure that existing profiles are updated. Once initiated, the import process for both modules is automated and requires no additional staff intervention (Figure 1).

The pre-processing of data prior to import requires approximately a half hour for each batch. Importing the user account data into Drupal requires up to a half hour if processing all student accounts. The initial import to create user accounts is completed prior to the start of the semester and includes over 10,000 accounts. Subsequent user account imports are much smaller as we only include newly registered students. The import of data into the user profiles requires more time and can take up to two hours for the initial load that occurs at the beginning of each semester. This process can run unmonitored once initiated.

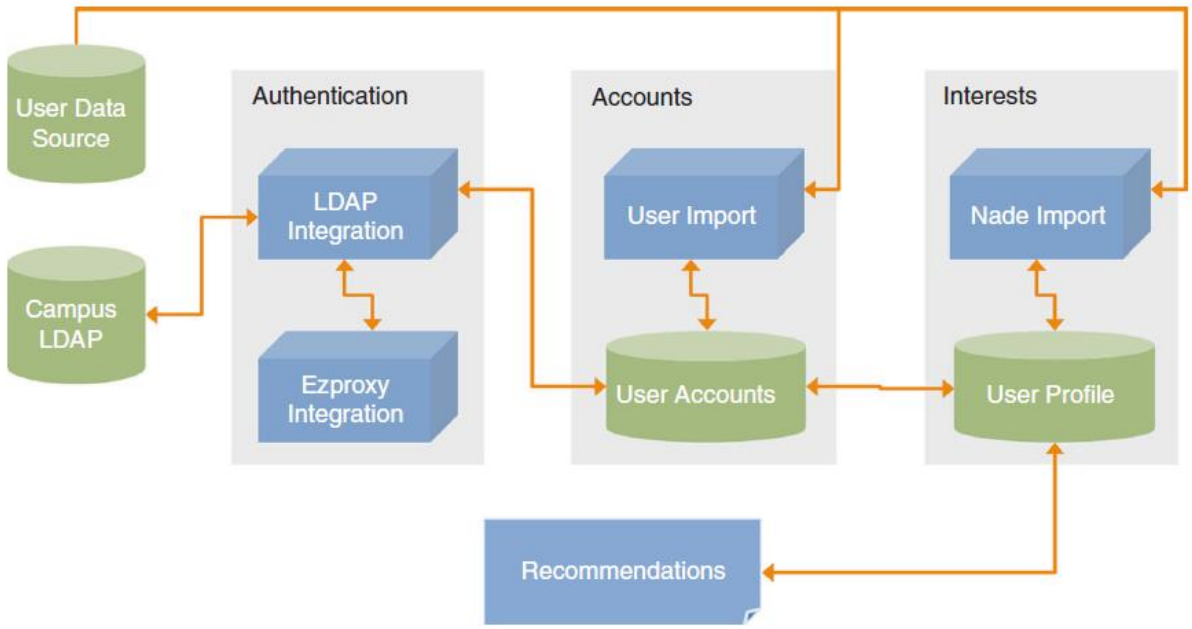

Figure 1.

Integration of user data and Drupal modules

When the import is complete, the system will generate a report that indicates if there were any errors. The number of imported profiles drops significantly for each subsequent data load as we only update profiles where the student has new or changed course enrollment information. By the fourth week of the semester, there are very few updates because students are no longer changing their course enrollments. By mid-semester, it is no longer necessary for us to run data imports as the enrolled courses for each student remains relatively static.

At the time of writing we are manually exporting the data from the standalone MySQL table and then using the import modules within Drupal to create or update user accounts. While that is not a lengthy or complicated process, we will examine the possibility of automating the data augmentation and transfer process. We also plan to significantly reduce the staff time required for import preparation by changing how our campus IT group provides the course enrollment data to us. One possible modification is to have the SIS data exported in a file that is separate from that which is provided to our ILS and formatted in a different manner. This would save us several steps in the data transformation process. Alternatively, we may look into having the SIS data exported directly into our standalone MySQL database. Adding and updating Drupal user data by querying the application programming interface of the ILS is another approach we will consider. We plan to upgrade Drupal to version 7 because its improved user account schema will merge the user account and user profile import processes.

Constructing lists of recommended content

As described in the preceding sections, the content relevance model is based on finding matches between the subjects associated with a student's course enrollments and the subjects associated with the content on

the library web site. In Drupal, the central taxonomy is the mechanism by which we cross-reference the subject in each user profile with the subjects of the content nodes.

To generate and display the lists of recommended resources that are derived from the content relevance model, we use the views add-on module for Drupal. This module "can fetch content from the database and present it to the user as lists, posts, galleries, tables, maps, graphs, menu items, blocks, reports, forum posts etc. It provides a graphical interface to a SQL query builder that can access virtually any information in your database and display it in any format" (Drupal, 2013). Once constructed, each set of queries is stored in an object referred to as a view. The output of a view can be placed anywhere within a Drupal web site. By not requiring expertise in the areas of programming and database queries, this module helps reduce the cost of building and maintaining the personalization system. This module also provides the means to filter output by variables within the content, taxonomy terms, user information, and node type attributes.

For the personalization-related views, it was necessary to apply query filters that were based on the taxonomy terms associated with a student's profile. This was accomplished through the use of the similar by terms module, an extension of the views module. As its name suggests, this module determines content relevance by examining taxonomy terms associated with each node in Drupal and then finding other nodes that share similar terms. The views created for the personalization system use this module to find library resource nodes that are relevant to the content profile node of each user.

In order to access the results of the personalization-related views and see their recommendations, users must login to Drupal. Authentication is therefore a critical component of the personalization system. The following section will describe the configuration of the Drupal authentication system in the context of the Library's existing systems and the personalization project.

Integrating the user authentication process

Two requirements for the personalization system were specifically relevant to the user authentication process: avoid disruption of the existing user experience and require no additional actions from the user. If an additional login was required to achieve personalization, this would have had a negative impact on the user's experience. This was especially true because there were already several separate logins required for various library web services. In our increasingly interconnected online world where authentication via Facebook and Google is readily available across many web sites, an additional login would have kept most of our users from trying the personalization features. Therefore, it was essential to integrate existing login mechanisms within the personalization system.

Using a combination of add-on modules and Drupal core functionality, we were able to create an integrated and seamless authentication process. This was primarily achieved through the use of two Drupal modules: LDAP integration and EZproxy. LDAP refers to a standardized protocol by which applications may interact with an organization's directory of personnel and account information. Many organizations provision LDAP servers as a centralized authentication and directory system for distributed web applications. Drupal's LDAP integration module provides a simple interface for setting-up a connection to an LDAP server. When a user attempts to sign-in on the CSUSM Library web site, this module checks with the campus LDAP system to verify the authenticity of the user's login information. If the user has an active campus account and is entering the correct username and password, the LDAP module will grant access to Drupal and update the user's basic account information. If the user has a valid campus account but is new to the Library's web site, the LDAP module will create a Drupal account for them. No library staff intervention is necessary as user account data and access rights are managed through the central campus system. For enhanced security, the authentication process is encrypted via the secure sockets layer protocol and the user's password is not saved to the Drupal database. This module's

ability to interact with the campus LDAP system ensures sustainable and efficient management of user accounts within the library web site.

In addition to integrating the Drupal user account system with the campus-wide authentication schema, the integration of pre-existing library authentication systems was an essential project requirement. At CSUSM, the most frequently accessed Library resources are licensed online databases. To grant access to these databases, the Library utilizes EZproxy (www.oclc.org/ezproxy.en.html). Initially developed by Chris Zager and now owned by OCLC, EZproxy is a ubiquitous application that allows libraries to easily grant their patrons authenticated access to licensed online content. EZproxy is typically configured to utilize an institution's authentication server, such as LDAP, or a library's ILS.

The Drupal EZproxy module integrates EZproxy authentication with the Drupal user account system and with the campus LDAP system. When a user's credentials are verified by Drupal, this module ensures that the user is automatically logged-in to EZproxy. This process is completely transparent and the user only logs in once to authenticate with both Drupal and EZproxy.

The LDAP and EZproxy modules were critical to the success of the personalization project. Without these Drupal add-ons, the Library would have had to develop its own modules. That would have made the project much more complex and would have greatly lengthened the time required for system development.

Generating personalized content for the user

The final step in building the personalization system was to create a visual framework for displaying personalized data on Library web pages. Essential requirements were to avoid a more complex interface and to reduce the number of steps required for accessing information resources. Users must be able to view personalized data within existing Library pages while also having access to a personalized profile.

A key component to achieving this outcome was the use of the panels module. The primary use for this module is to pull together content and data from any area within a Drupal site and to display that information as a grouping of panels within a single page. Panels is an integral element of our personalization system because it passes user account data to the views database queries, thus filtering the results to that which is relevant to the user.

An example of the visual display generated by the panels modules is shown in Figure 2. In that display, the recommendations tab of a user's profile page is configured to display three different views, each in its own panel: course guides, databases, and research guides. The content of each of view is derived from the user data passed to it via the panels module. This module also ensures that each student is only able to access their own profile and recommendations.

Personalizing the library home page

Analysis of the Library web site's usage data has shown that a vast majority of our users start at the Library homepage. Adding personalized results to the homepage would promote the visibility of recommended resources and simplify access. Users would not have to navigate to a different page in order to view their personalized recommendations.

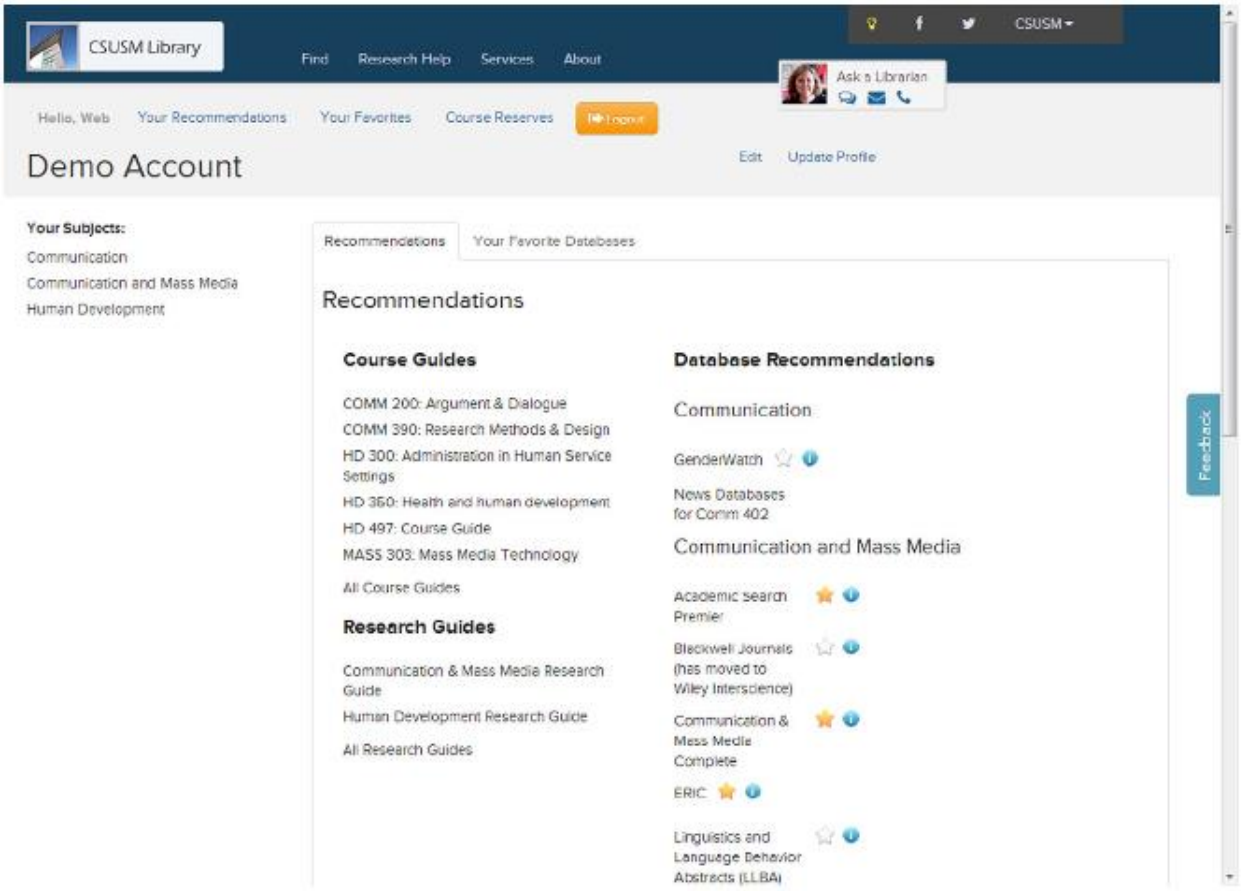

Figure 2.

A personalized user profile page

If a visitor to the Library web site successfully authenticates as a CSUSM student, the panels module will rearrange the homepage to prominently display personalized content panels (Figure 3). As shown in Figure 3, the right column adds a list of recommended research guides and displays the pictures and contact information for selected subject-specialist librarians. The center column of the page offers recommendations for subject-specific database lists. If the user is not logged-in, the module will display the default Library homepage layout and content.

Personalized user menu

To provide access to the user's profile page, we adopted commonly used terminology and navigational aids employed by e-commerce sites that offer personalization. On Amazon.com, users who have logged-in to their account have the option to access "Your Amazon.com." Once the user navigates to that area of Amazon, the site makes available a horizontal list of personalization-oriented menu items. The menu options include "Recommended for You" and "Your Profile." This is a convenient way to visually highlight the recommendations for the logged-in user. Netflix and Goodreads, also place the

personalization links into a horizontal menu below the site header. The Library's web site emulates those user experience conventions by also providing links to recommendation pages in a horizontal menu directly below the site header.

Project assessment

Several assessment methods for determining the effectiveness of personalization are in-use or planned. One methodology is to collect usage data via web analytics tools. The applications we are using for this are Google Analytics (www.google.com/analytics),

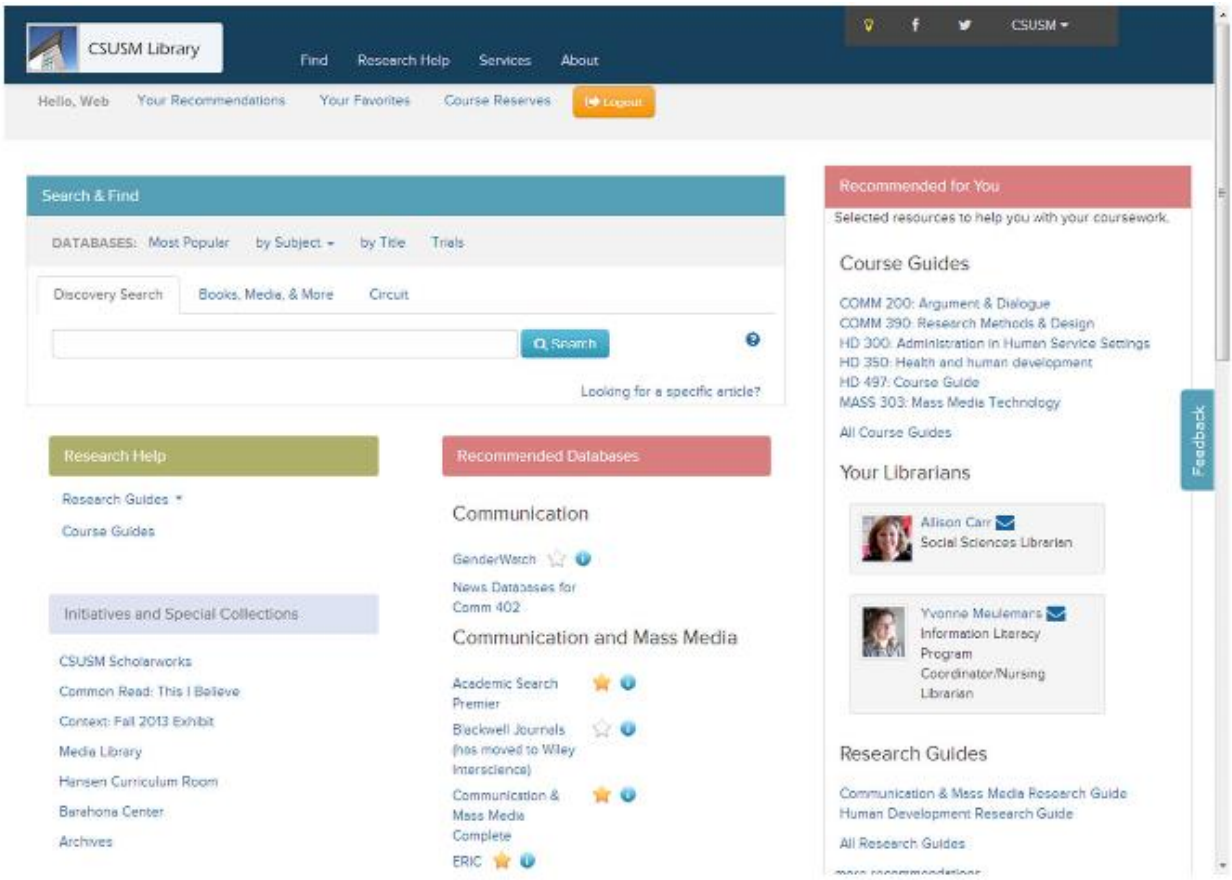

Figure 3.

Personalized library home page

Piwik (http://piwik.org), and CrazyEgg (www.crazyegg.com). Google Analytics is a free, hosted service that captures usage data via a small JavaScript snippet that is embedded within the source code of each Library web page. The usage data available through this service includes the number of times a page is viewed, user visits, visitor characteristics, and navigational paths. While Piwik is very similar to Google Analytics, it organizes and stores its collected data in a different manner. Together, these tools present information that is helpful to understanding a web site's usage patterns. The analytics data gathered by those applications will show whether users clicked on links provided via the personalization system.

For this project, the transitions data in Piwik (Figure 4) and the navigation summaries in Google Analytics are especially useful. As seen in Figure 4, a Piwik transitions map can show whether a user arrived at a page via a personalized link. In the example given, the visits to the BUS302 course guide page from https://biblio.csusm.edu and/research/dashboard are via personalized links. This is because the Library homepage in its non-personalized state does not offer direct links to individual course guides.

The Library also utilizes an internal database that collects data on the number of times users click on e-resources within the Library web site. This is useful for measuring when and where our users elect to visit Library e-resources. This database does not track the number of searches within an e-resource. The data gathered is used to analyze the effectiveness of link placement and the degree to which users access specific e-resources from the Library web site. For the personalization project, this method of click-through tracking is offers data on whether users clicked on personalized e-resource links.

In contrast to aforementioned assessment tools, CrazyEgg focusses on click-throughs within a single web page. Its heatmap report "is a visualization of where your visitors are clicking. The brighter the area, the more popular it is. The darker the area, the less popular it is" (Crazy Egg, n.d.). In addition, it can offer a tally of clicks per link for a specific web page (Figure 5). CrazyEgg is an effective tool for measuring the effectiveness of link placement, labels, and content within specific web pages.

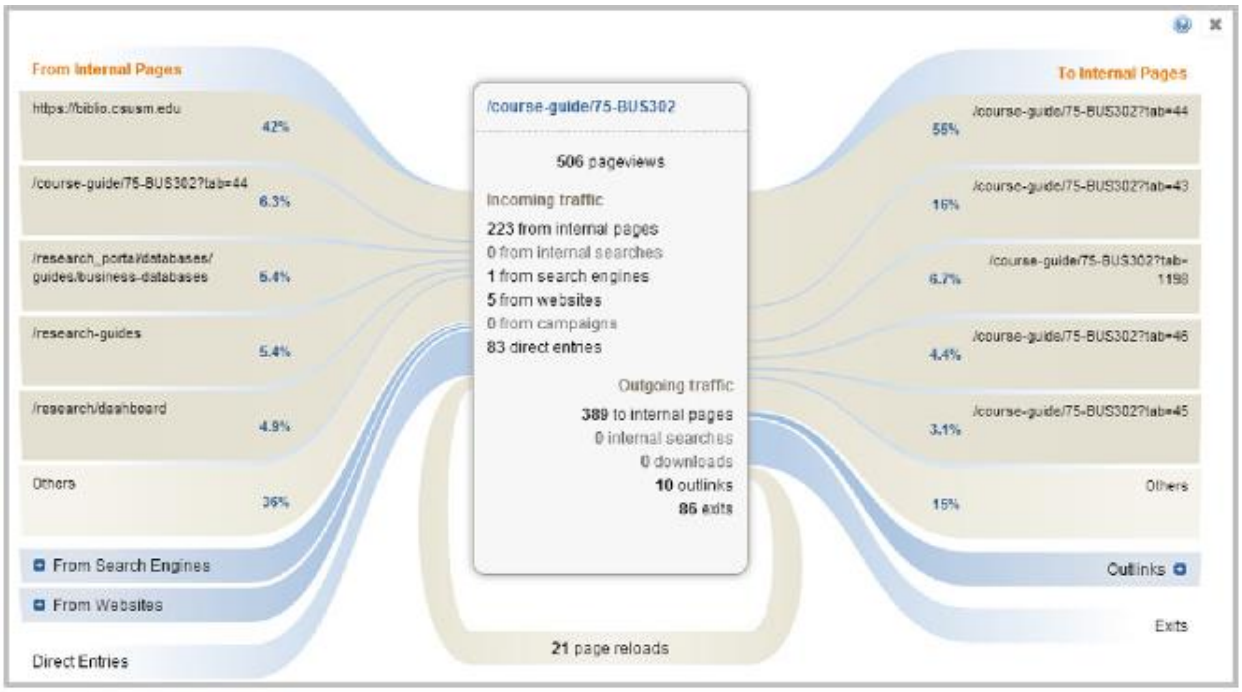

Figure 4.

Piwik transitions map

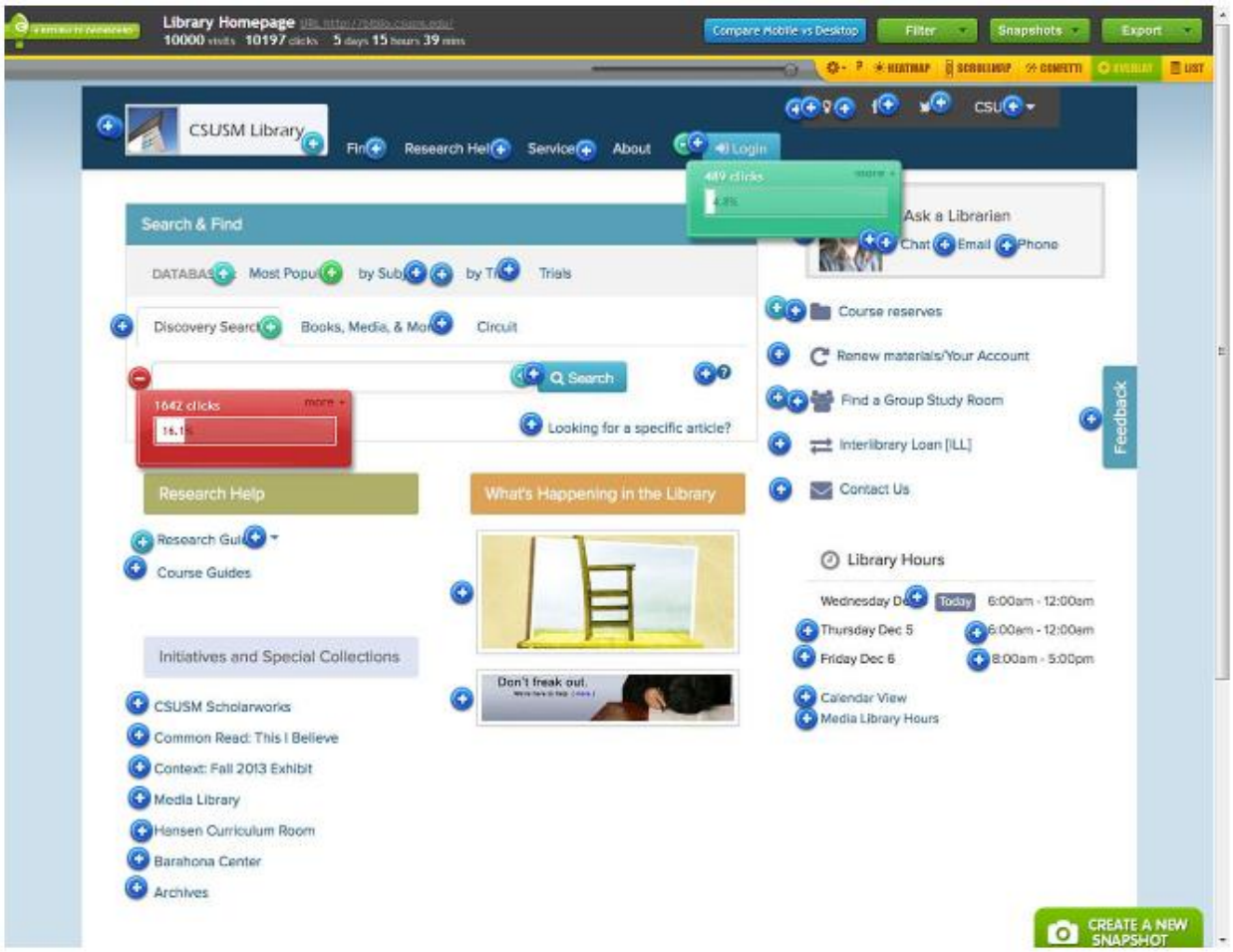

Figure 5. CrazyEgg overlay of clicked links

These usage data gathering tools will help answer a number of questions as we assess the outcomes of this project. Are visitors using the specific web pages associated with the personalization features? Do visitors use the new pathways made available by personalization to access library resources? Is the user interface of the personalized homepage effective? By gathering data to answer these questions, the Library can begin to measure the success of personalization, focus on areas needing improvement, and generate ideas for building new features.

Preliminary analysis of the usage data collected via the tools listed above indicates that students are using the personalized links on the Library home page. One basis for this assessment is a CrazyEgg usage test conducted between December 4, 2013 and December 10, 2013. This test tracked clicks on the Library home page over the course of 10,000 visits. To assess if personalized links where used, we examined the test results for links that were only visible after a user had logged-in. One link examined was that which is used to access JSTOR. We used this because it is becomes visible on the Library home page only after users login and if their personalized Library home page includes JSTOR as a recommendation. The test results included 105 clicks on personalized JSTOR links, from the Library home page. During the test period, all JSTOR links, from anywhere on the Library web site, were clicked 339 times. Based on that data, 30 percent of the click-throughs to JSTOR, from the Library web site, were

via personalized links. It is important to note that a minimum of three clicks are required to access the JSTOR links when starting from the default, non-personalized Library homepage.

To perform a preliminary analysis of personalized course guide links, we utilized data collected via Google Analytics. The focus of this analysis was on a selection of the more heavily used course guides. The time span of the data analyzed was from November 11, 2013 through December 9, 2013. To be considered a visit from a personalized link, the page from which a user arrived at a course guide was required to be either the library home page or the user's profile page. A preliminary analysis of the data is presented in Table I.

These initial findings indicate that personalized links are being used. Use of these links has the effect of reducing the number of clicks required to access library content. In addition, they highlight resources that may not have been known to students, especially those who are new to the library. As such, these early results indicate that the personalization system is beneficial to our users.

We will need to conduct much more in-depth analysis of our data to better understand the impact of personalization. Testing, fine-tuning, and expanding our data-gathering methodologies are essential steps toward ensuring adequate assessment of the system. For example, personalized pages are not consistently identified within our system. As a result, it is possible that some clicks from personalized links are not being counted within our analysis. Usage of course guides fluctuates widely over the course of an academic semester. The use of course guides can vary greatly among academic departments. Therefore, we should correlate usage data with differences in library web site usage by subject discipline and by time of year. In-person and online surveys would provide an important complement to the data collected via our online analytics tools. These surveys would measure user satisfaction and gather insight on potential improvements to the personalization system.

Future directions

The next phase of the personalization project will expand its scope to include a number of enhancements. One of these enhancements will add personalized elements that focus on specific user groups such as faculty, graduate students, and extended learning students. Faculty feedback has indicated that they would like to have easier access to the web site tools that they use most. This might include personalizing the homepage by adding links to forms that are for faculty-use only. Another use of faculty-specific personalization is to display the name and contact information for their subject librarian and to provide a list of new titles specific to their disciplines and research interests. Providing course-specific content within the campus learning management system is another future goal.

Conclusion

For academic libraries, student course enrollment data is a vast source of existing user information that offers great potential for the personalization of online systems. Libraries and their parent institutions have made very limited efforts toward developing systems that integrate this user data to build intelligent web sites.

Through the use of adaptive personalization, the CSUSM Library web site now generates a customized user profile based on a student's course enrollments. Using a subject-specific taxonomy, those course selections are dynamically cross-referenced with relevant Library resources. This allows the Library to provide tailored recommendations that are specific to a student's research needs, such as research

Course guide	Pageviews from personalized links	Total pageviews site-wide	% from personalized links
/guides/course-guide/75-BUS302	202	523	39
/guides/course-guide/116-BIOL353	81	110	74
/guides/course-guide/345-BA650	21	52	40
/guides/course-guide/343-LTWR115	34	141	24
/guides/course-guide/256-HIST301	35	62	56
/guides/course-guide/78-GEO102	26	85	31

Table 1. Pageviews from personalized course guide links as percentage of total pageviews, Novermber 11, 2013-December 9, 2013

guides and subject librarians. The Library was able to accomplish this largely through the use of open source web applications and a limited amount of custom programming.

The adaptive personalization system developed for this project requires a relatively low-resource commitment on the part of the library and its parent institution. It enriches the user experience while requiring no additional steps from students. A personalization project of this nature is an investment toward developing more user-centric library web sites and online services. This project brings into focus and opens the door for further investigation into the potential benefits of leveraging course enrollment data to enhance the delivery of library information resources and instruction.

References

Brantley, S., Armstrong, A. and Lewis, K.M. (2006), "Usability testing of a customizable library web portal", College & Research Libraries, Vol. 67 No. 2, pp. 146-163.

Ciccone, K. (2005), "MyLibrary@NCState – a library portal after five years", Journal of Library Administration, Vol. 43 No. 1, pp. 19-35.

Cohen, S., Fereira, J., Horne, A., Kibbee, B., Mistlebauer, H. and Smith, A. (2000), "MyLibrary: personalized electronic services in the Cornell University Library", D-Lib Magazine, Vol. 6 No. 4, available at: www.dlib.org/dlib/april00/mistlebauer/04mistlebauer.html (accessed September 23, 2012).

Crazy Egg (n.d.), "Crazy egg – training center – heatmaps", Crazy Egg Inc., available at: www.crazyegg.com/training_center/tutorial/Heatmaps (accessed May 20, 2013).

Drupal (2013), "Working with views", Drupal.org, available at: https://drupal.org/documentation/modules/views (accessed February 4, 2014).

Eirinaki, M. and Vazirgiannis, M. (2003), "Web mining for web personalization", ACM Transactions on Internet Technology, Vol. 3 No. 1, pp. 1-27.

Forsati, R. and Meybodi, M.R. (2010), "Effective page recommendation algorithms based on distributed learning automata and weighted association rules", Expert Systems with Applications, Vol. 37 No. 2, pp. 1316-1330.

Frias-Martinez, E., Chen, S.Y. and Liu, X. (2009), "Evaluation of a personalized digital library based on cognitive styles: adaptivity vs adaptability", International Journal of Information Management, Vol. 29 No. 1, pp. 48-56.

Frias-Martinez, E., Magoulas, G., Chen, S. and Macredie, R. (2006), "Automated user modeling for personalized digital libraries", International Journal of Information Management, Vol. 26 No. 3, pp. 234-248.

Ghaphery, J. (2002), "My library at Virginia Commonwealth University", D-Lib Magazine, Vol. 8 Nos 7/8, available at: www.dlib.org/dlib/july02/ghaphery/07ghaphery.html (accessed September 23, 2012).

Gibbons, S. (2003), "Building upon the MyLibrary concept to better meet the information needs of college students", D-Lib Magazine, Vol. 9 No. 3, available at: www.dlib.org/dlib/march03/gibbons/03gibbons.html (accessed September 23, 2012).

Heinrichs, J.H., Lim, K., Lim, J. and Spangenberg, M.A. (2007), "Determining factors of academic library web site usage", Journal of the American Society for Information Science and Technology, Vol. 58 No. 14, pp. 2325-2334.

Jeevan, V.K.J. and Padhi, P. (2006), "A selective review of research in content personalization", Library Review, Vol. 55 No. 9, pp. 556-586.

Ketchell, D.S. (2000), "Too many channels: making sense out of portals and personalization", Information Technology and Libraries, Vol. 19 No. 4, pp. 175-179.

Kim, Y. (2010), "The adoption of university library web site resources: a multigroup analysis", Journal of the American Society for Information Science and Technology, Vol. 61 No. 5, pp. 978-993.

Kim, Y.-M. (2011), "Users' perceptions of university library websites: a unifying view", Library & Information Science Research, Vol. 33 No. 1, pp. 63-72.

Kumar, N. and Benbasat, I. (2006), "The influence of recommendations and consumer reviews on evaluations of websites", Information Systems Research, Vol. 17 No. 4, pp. 425-441.

Lee, C.H. and Cranage, D.A. (2011), "Personalization – privacy paradox: the effects of personalization and privacy assurance on customer responses to travel web sites", Tourism Management, Vol. 32 No. 5, pp. 987-994.

Liang, T.-P., Lai, H.-J. and Ku, Y.-C. (2007), "Personalized content recommendation and user satisfaction: theoretical synthesis and empirical findings", Journal of Management Information Systems, Vol. 23 No. 3, pp. 45-70.

Liu, S. (2008), "Engaging users: the future of academic library web sites", College & Research Libraries, Vol. 69 No. 1, pp. 6-27.

Morgan, E.L. (1999), "MyLibrary: a model for implementing a user-centered, customizable interface to a library's collection of information resources", arXiv:cs/9902003, available at: http://arxiv.org/abs/cs/9902003 (accessed September 14, 2011).

Munro, K. and McLure, M. (2010), "Research for design: exploring student and instructor attitudes toward accessing library resources and services from course management systems (CMS)", Communications in Information Literacy, Vol. 4 No. 1, pp. 29-60.

Nichols, J. and Mellinger, M. (2007), "Portals for undergraduate subject searching: are they worth it?", Portal: Libraries and the Academy, Vol. 7 No. 4, pp. 481-490.

Park, D.H., Kim, H.K., Choi, I.Y. and Kim, J.K. (2012), "A literature review and classification of recommender systems research", Expert Systems with Applications, Vol. 39 No. 11, pp. 10059-10072.

Porter, B. (2011), "Millennial undergraduate research strategies in web and library information retrieval systems", Journal of Web Librarianship, Vol. 5 No. 4, pp. 267-285.

Pu, P., Chen, L. and Hu, R. (2012), "Evaluating recommender systems from the user's perspective: survey of the state of the art", User Modeling and User-Adapted Interaction, Vol. 22 Nos 4/5, pp. 317-355.

Shedlock, J., Frisque, M., Hunt, S., Walton, L., Handler, J. and Gillam, M. (2010), "Case study: the Health SmartLibrary experiences in web personalization and customization at the Galter health sciences library, Northwestern University", Journal of the Medical Library Association, Vol. 98 No. 2, pp. 98-104.

Sirisha, B.S., Jeevan, V.K.J., Kumar, R.V.R. and Goswami, A. (2009), "A personalized information support system for searching portals and e-resources", Program: Electronic Library and Information Systems, Vol. 43 No. 1, pp. 77-93.

Smeaton, A.F. and Callan, J. (2005), "Personalization and recommender systems in digital libraries", International Journal on Digital Libraries, Vol. 5 No. 4, pp. 299-308.

Somerville, M.M. and Brar, N. (2009), "A user-centered and evidence-based approach for digital library projects", The Electronic Library, Vol. 27 No. 3, pp. 409-425.

Sunikka, A. and Bragge, J. (2012), "Applying text-mining to personalization and customization research literature – who, what and where?", Expert Systems with Applications, Vol. 39 No. 11, pp. 10049-10058.

Wang, M. and Yen, B. (2010), "The effects of website personalization on user intention to return through cognitive beliefs and affective reactions", paper presented at Pacific Asia Conference on Information Systems, July 9-12, Taipei, available at: http://aisel.aisnet.org/pacis2010/155 (accessed May 20, 2013).

Further reading

Biddix, J.P., Chung, C.J. and Park, H.W. (2011), "Convenience or credibility? A study of college student online research behaviors", The Internet and Higher Education, Vol. 14 No. 3, pp. 175-182.

Chen, C.-M., Lee, H.-M. and Chen, Y.-H. (2005), "Personalized e-learning system using item response theory", Computers & Education, Vol. 44 No. 3, pp. 237-255.

De Rosa, C., Gauder, B., Cellentani, D., Dalrymple, T. and Olszewski, L.J. (2011), "Perceptions of libraries, 2010: context and community: a report to the OCLC membership, OCLC", available at: www.oclc.org/content/dam/oclc/reports/2010perceptions/2010perceptions_all.pdf (accessed November 8, 2012).

Jafari, A., McGee, P. and Carmean, C. (2006), "Managing courses defining learning: what faculty, students, and administrators want", EDUCAUSE Review, Vol. 41 No. 4, pp. 50-52, 54, 56-58, 60, 62, 64, 66, 68, 70.

Jeong, H.-Y., Choi, C.-R. and Song, Y.-J. (2012), "Personalized learning course planner with e-learning DSS using user profile", Expert Systems with Applications, Vol. 39 No. 3, pp. 2567-2577.

Morgan, E.L. (2002), "Making information easier to find with MyLibrary", Musings, available at: http://infomotions.com/musings/mylibrary-access/ (accessed September 14, 2012).

Morgan, K. and Reade, T. (2000), "Pioneering portals: MyLibrary@NCState", Information Technology and Libraries, Vol. 19 No. 4, pp. 191-198.

Mulvenna, M.D., Anand, S.S. and Bu¨chner, A.G. (2000), "Personalization on the net using web mining: introduction", Communications of the ACM, Vol. 43 No. 8, pp. 122-125.

Vaughan, J. (2011), Web Scale Discovery Services, American Library Association, Chicago, IL.

Zhou, J. (2003), "A history of web portals and their development in libraries", Information Technology and Libraries, Vol. 22 No. 3, pp. 119-128.

About the author

Ian Chan currently serves as the Systems Coordinator and Web Development Librarian at the California State University San Marcos. He leads and implements initiatives that focus on integration of new

technology in the library and improving the user's research experience. At present, his projects include enhancing personalization within the library's web site, integrating discovery search, and providing course-specific library data to the campus learning management system. Ian Chan can be contacted at: ichan@csusm.edu