User-Centric Solutions for Scholarly Research in the Library

by TAMAR SADEH

ABSTRACT

In recent years, the behavior of library users has been undergoing change, as they rely increasingly on web search engines and internet services, rather than libraries, to fulfill their information needs. This paper describes the expectations of such users when looking for information and presents an optimized approach that enables research libraries to provide their patrons with an up-to-date user experience while capitalizing on the unique advantages offered by libraries. Ways of integrating the new solution with the current library environment are discussed, as well as future directions. The Primo discovery and delivery system from Ex Libris serves as an example of such solutions.

INTRODUCTION

The publication of the OCLC Online Computer Library Center survey Perceptions of Libraries and Information Resources in late 2005, followed by a report focusing only on the student participants in the survey - College Students’ Perceptions of Libraries and Information Resources - confirmed what libraries already knew: the behavior of library users has changed since the emergence of web search engines, online bookstores, and other services available on the Web that enable many users to satisfy their information needs without the help of the library (OCLC, 2005; OCLC 2006). According to the OCLC reports, 89% of the undergraduate and graduate students surveyed start their search for information with web search engines, and only 2% start at the library web site. Other statistical data support this trend, even though the picture portrayed by the data is not as grave. For example, statistics compiled by the Association of Research Libraries (ARL), whose membership spans North America, indicate a significant decrease in reference transactions (33%) and circulation transactions (6%) at ARL institutions between 2000 and 2006, despite an increase of about 12% in the number of students enrolled in these institutions at that time.1

Looking for ways to retain their users and maintain their hegemony as information providers, libraries are considering new approaches - user-centric solutions to replace their traditional online user interfaces. This paper discusses the expectations of today’s users and how these expectations can be addressed in the context of research libraries.

USERS’ PREFERENCES

The attractiveness of web search engines is clear: 84% of the student participants in the OCLC survey find web search engines much more convenient than online or physical libraries; 87% find the web search engines easier to use; and 90% consider them faster. The students also feel that web search engines are more cost-effective (71%) and even more reliable (63%) than libraries (OCLC, 2006). However, behavior patterns that one can readily observe in the broader scope of user activities reveal additional reasons for which people prefer resources that are not part of the library offering, even for scholarly research.

A detailed discussion of user preferences is described elsewhere (Sadeh, 2007; Sadeh, in press). To summarize, we can identify the following characteristics of web search engines and other internet services that are attractive to end users:

- Simple searching: Web search engines and internet services do not require a significant degree of searching literacy on the part of the user and provide results that are usually good enough without precision searching. Features such as spelling corrections and suggested alternatives to search queries that yield few results
relieve the user from the burden of entering highly accurate search terms. Web search interfaces are typically very simple and require no expertise for conducting a keyword search - the kind of search that is most popular on the web. Furthermore, recent improvements in the way in which results are organized and displayed on the screen have introduced ‘faceted browsing’, a method by which users can easily narrow down their searches to obtain more precise results.

Vast, heterogeneous content: The content coverage and heterogeneity of resources offered to end users by web search engines and internet services such as Google Scholar and Windows Live Search are considerably greater than what libraries offer. Although the quality of the content provided through web searches may fall short of that provided by the carefully selected library collections, users are attracted to the ease of one-stop shopping for all types of materials. Also, such users perceive searching in library collections, which are usually dispersed among multiple systems with diverse interfaces - a catalog, digital repositories, remote databases, and more - as a complex, challenging process.

Ease of access: Whereas library systems adhere to the traditional framework of the web and offer distinct entry points to their collections, web search engines and internet services have developed alternative entry points that enable users to access such services anytime and anywhere. For example, users can conduct a search through the Google toolbar in their browser or through a Google desktop gadget, and save any web page to their Connotea account. Furthermore, web search engines and internet services are always available for discovery and pose no obstacles, such as requiring users to connect through a proxy server or to sign in before beginning to search.

Online availability of material: Most users prefer material that is available online and that they can access from afar and use in ways that are suited to their research style and tempo. Online access facilitates searching within documents, homing in on a specific paragraph, zooming into and out of images, watching videos, listening to audio items, and extracting quotations. Although libraries already provide large quantities of materials in electronic form, many users still perceive libraries as a place for physical items and often turn to the broader web to find electronic materials.

Interactive experience: When comparing the overall user experience provided by library systems to that provided by popular internet services such as Amazon.com, del.icio.us, Facebook, Flickr, MySpace, and YouTube, we see that the basis of the latter group is an interactive platform through which users not only find information but also annotate material, write reviews, carry on discussions, and even make material that they have created available to other community members. Library systems, at this point in time, typically serve only as sources of information; the characteristics of Web 2.0 as a platform for collaboration are only now emerging in such systems.

Recommendations: Gathering information regarding users’ preferences - explicitly, by providing tools with which people can rate and review materials, or implicitly, by logging user behavior - enables internet services to present recommendations that help users determine the value of specific items for their work. For example, a user whose search in Amazon turns up several similar books, can read the reviews added by other community members to figure out which book will be most helpful. In the library context, where many users have no direct contact with a librarian but rather access the library collections from afar, few such mechanisms are evident.

If, in the past, searching for information occurred primarily in the scholarly domain and users’ information-retrieval behavior affected how they searched for non-scholarly information, the researchers of today arrive at institutions of higher education already accustomed to the internet environment and derive their expectations from it. In a report commissioned by the Bibliographic Services Task Force of the University of California, the authors conclude that their ‘users expect simplicity and immediate reward and Amazon, Google, and iTunes are the standards against which we are judged. Our current systems pale beside them’ (University, 2005). The challenge for libraries, therefore, is to determine how they wish to portray themselves to their users and how they can best serve the institutions to which they belong. In this era of shrinking library budgets and the changes in business models, the accountability of libraries regarding the resources that they create, acquire, and license, on the one hand, and the usage of their collections, on the other hand, may prove to be crucial for their future.
THE ROLE OF THE RESEARCH LIBRARY TODAY

Before looking at technological solutions for addressing the expectations of today’s users, let’s examine the role of the research library in our era, in which libraries are no longer the only source of information for scholars. In her lecture at the September 2006 meeting of the International Group of Ex Libris Users, Karen Calhoun envisioned a “new kind of library”, one that is “... more involved with research and learning materials and systems ... more engaged with campus communities”, a library that “make[s] library collections and librarians more visible” and “move[s] to next generation systems and services.” Calhoun suggested that “the continued success of the research library is linked to the integration of the library into the larger information network processes of the university and indeed, of the world at large” (Calhoun, 2006b).

When designing ‘next generation systems and services’, libraries need to build on their unique qualities and to complement what web search engines and internet services can offer. There is no doubt that in the future research environment, various systems and services will coexist; libraries are thus charged with finding the framework in which their contribution to the research process will be as meaningful as possible. Indeed, to address the exact research needs of their users, libraries can capitalize on the qualities that differentiate them from web search engines and Internet services:

- **Collections**: Despite the rapid pace of digitization, especially that of books, by major players such as Google, Yahoo, and Microsoft in collaboration with libraries, physical collections will continue to be crucial for research in many disciplines, at least for the time being. Moreover, numerous libraries hold special collections - physical or digital - that are not available elsewhere. In the future, libraries might also consider serving as publishers of locally held research materials that are not made available to researchers today, such as working papers and data sets.

- **Quality and focus**: Unlike web search engines and internet services, research libraries select the information to be made available: libraries acquire and license only materials that are deemed relevant to and of sufficient quality for the specific library community. In a research environment, quality measures are crucial, and library users are aware of the selective nature of the library collections. For example, more than 75% of the students surveyed by OCLC perceive libraries as a more “accurate” and “trustworthy or credible source” than web search engines (OCLC, 2006). In addition, libraries can create subsets of collections and can group collections to form search scopes that are most likely to address the needs of a specific group of researchers, thus helping researchers focus on the materials that are most relevant to them instead of spending time sifting through materials of questionable value.

- **Metadata**: Although web search engines’ reliance on free-text indexing yields remarkably good results, the metadata available for library collections is invaluable, and sometimes -when items are not textual or are not available online - irreplaceable. As current initiatives and discussions indicate, cataloging practices may well change in the future; meanwhile, however, the metadata that is available to libraries enables them to describe their collections in greater detail, thus providing for a better search experience, one that is more applicable to their communities.

- **User context**: When it comes to context-specific services, libraries find themselves in a unique position. Because they know their own communities, libraries can tailor the services to a specific user, taking into account not only the profile and preference of that user but also the library’s subscriptions and policies. Hence, even though users can discover materials on the web by means of non-library interfaces, they will often turn to the library for obtaining these materials. And by monitoring the usage of available materials and the requests for materials that are not available, libraries can develop their collections to better address their users’ needs.

- **Branding**: Although libraries are typically open, at least partially, to whoever wishes to search their collections, their main goal is to serve specific communities. To gain visibility on the part of their users, on the one hand, and the institutions to which the libraries belong, on the other hand, libraries need to market themselves effectively; among other things, they need to create their own distinct branding. A library can create a look and feel characterizing its services on its web site and elsewhere. A good example of a branded component is a link resolver: a library designs a button that indicates the availability of relevant services, assigns a name to it (such as More Info or Full Text), and makes this button available in multiple resources in the users’ research environment, such as the library catalog, abstracting and indexing databases, and Google Scholar. When a user clicks the button, the system displays a default service (typically the full text of an article) or a menu of relevant services from which the user can select the one that is the most
appropriate. Both the button and the menu of services carry the library’s branding, presumably creating the desired impact on users. Branding refers not only to the look and feel of the services but also to terminology and interface languages.

- **Collaboration:** Libraries have a long tradition of collaboration with each other. Collaboration becomes even more important in the era of globalization and changing economies, and when supported by appropriate technologies, it can open up new frameworks for sharing collections and services. Collaboration extends the reach of a library and enables it to offer better and more accurate services to its users. Current technological trends based on interoperability and the use of standards are setting the ground for cooperation among libraries and providers of content and services; together, these bodies will create an environment that best suits the needs of twenty-first century researchers.

Providing quality materials to their users is no doubt one of the major tasks of libraries, particularly research libraries. However, while examining solutions that can satisfy their users’ needs, libraries should be looking at the larger context of their activities and moving forward in a manner that capitalizes on past investments; that is cost-effective regarding development, implementation, and maintenance; and that will enable the libraries to meet future needs in terms of technological infrastructure, scalability, sustainability, and adaptability to expected and unexpected change.

**NEXT-GENERATION LIBRARY SYSTEMS**

Since the launch of the new library catalog at North Carolina State University in January 2006 and the publication of several major reports - the OCLC survey, the University of California Bibliographic Services Task Force report, and a report prepared for the Library of Congress - the ‘next-generation library catalog’ has become the object of many discussions at conferences and in online forums, and the topic of numerous publications (OCLC 2005; OCLC 2006; University, 2005; Calhoun, 2006a; Antelman et al., 2006; Casey & Savastinuk, 2006; Morgan, 2007). As pointed out by Marshall Breeding, the phrase ‘next-generation library catalog’ does not accurately reflect the kind of systems that have emerged in the last few years, because these systems are intended to be more than just a replacement for the traditional library catalog (Breeding, 2007). And because some of them have already been implemented, the adjective ‘next-generation’ is misleading. However, regardless of the term used, the new systems are meant to help libraries address the expectations of today’s users.

Although these new systems differ from each other in some ways, they are all based on decoupled architecture - the separation of the user experience from the management of the library collections. With the advent of decoupled architecture, a library’s current systems, such as an integrated library system and digital asset management system that have been tailored to the needs of the librarians, can continue to fulfill administrative functions while the interface for users is being developed as a new, separate platform designed around their needs. Such a framework can provide libraries with several benefits:

- **Flexibility:** Because it is not tied to the library’s management systems, the user experience suits the needs of the end users. The user experience is free of constraints imposed by management systems in terms of specific data structures and administrative workflows.

- **Unification:** Data coming from heterogeneous, harvestable resources can form one source of information for end users. The fact that data elements originate from separate sources becomes irrelevant to the discovery process.

- **Data enhancement:** Once it harvests the data from the administrative systems, the new system can easily enhance that data with additional data elements from other, typically third-party resources. Such data elements include images of book covers, tables of contents, descriptions, and reviews. Furthermore, more heterogeneous data resources can be added in the future.

- **Integration:** Adherence to the newest standards of interoperability enables the library to easily integrate its new interface into an existing environment and to make the interface interoperable with various systems in the user’s work space, such as course management systems and reference management tools.

- **Step-by-step evolution:** With a system deploying decoupled architecture, libraries can offer a new, improved user experience that addresses their users’ expectations without disturbing the behind-the-scenes
TAMAR SADEH

administrative systems. Similarly, libraries can add, modify, and replace components of their administrative systems without affecting the user experience.

Decoupled architecture lays the groundwork for a robust service for end users, on the one hand, and a unified resource management system, on the other hand. The first already exists. The second will be the challenge ahead.

Several vendors have developed or are in the process of developing a user-experience solution based on decoupled architecture. Among these are Endeca, a company that originated in the e-commerce arena and adapted its search engine for bibliographic data searches; Ex Libris, with its Primo discovery and delivery solution; Innovative Interfaces, developer of the Encore discovery services platform; and Medialab Solutions (recently acquired by Bowker), with its AquaBrowser Library product. Other initiatives include OCLC WorldCat Local, which is a version of the OCLC WorldCat global library network that can be tailored, to some degree, to individual libraries. A detailed description of the various solutions is available in a recent report by Marshall Breeding (Breeding, 2007).

The next sections of this paper provide a closer look at Primo as an example of a new solution for addressing the expectations of today’s users. The design concepts of Primo are described in greater detail in a case study by Sadeh (Sadeh, in press).

PRIMO DESIGN CONCEPTS

Decoupled architecture: what does it mean?

Adhering to the principle of decoupled architecture, Primo includes a publishing platform that enables it to harvest data from resources controlled by the library, such as the library catalog, digital repositories, knowledge bases that describe the library’s electronic holdings, and course management systems. This data is typically bibliographic data representing items of any type that the library offers - for example, books, journals, articles, images, music scores, videos, and audio - and may also include textual objects, such as PDF files stored in a local repository. The publishing platform normalizes the harvested data (that is, makes the data conform to one set of rules), converts all the data to one format, and enriches it with information such as book-cover images, abstracts, and tables of contents from third-party sources such as Amazon.com and Syndetic Solutions. In addition, the publishing platform detects duplicate records and groups all the similar records (for example, multiple editions of the same book) to display them to the user as one entity in the initial search results. The outcome of this process is the Primo index, which is optimized for Primo searches (see fig. 1). After the initial building of the index, the publishing platform also maintains it; whenever the contents of the source repository change, the publishing platform harvests the data that was added or changed and takes it through the publishing process.

Figure 1. Decoupled architecture
The interaction between Primo and the administrative, or back-office, systems is bidirectional: on the one hand, Primo harvests data from the back-office systems to build and maintain its index; on the other hand, it connects the user to the delivery services that are provided or initiated by the systems containing the data. Such services can include detailed holdings of physical library materials, a PDF reader for full-text articles, a viewer for images, and a player for video or audio files.

As described earlier, the Primo index includes only harvestable data. However, a typical academic research environment consists of multiple remote databases that are usually not harvestable; often license or technology constraints prevent harvesting, or the library does not wish to replicate and maintain data that is stored elsewhere. Through the application programming interface of the Ex Libris MetaLib metasearch system, Primo employs a just-in-time approach for accessing such remote resources; that is, instead of harvesting data from remote resources on a regular basis, the Primo system searches such resources when a user submits a query. Despite the different technologies involved in searching in local resources and remote resources, the user, for whom the location of materials is irrelevant, sees only one interface and one starting point for searching for all materials.

User experience

The developers of Primo identified the major issues with which typical users struggle and addressed these issues in the solution’s user-centered design.

Where should I search?
When implementing Primo, a library determines the overall scope provided for users’ searches by selecting the resources that it harvests and the remote materials that it offers through metasearching. Furthermore, a library can establish multiple search scopes and tailor them to the needs of specific user groups. For example, in a consortium environment, a specific library can define only its own collections as the default search scope, but the user can choose to look for materials originating in any or all of the consortium members’ libraries; alternatively, the library can set the collections of all the member libraries as the default scope, but the user can also choose the scope of his or her own institution.

How do I search?
Because it is based on paradigms that are well known to end users, the Primo search interface requires no training, as was confirmed in usability studies conducted by the University of Minnesota (Rosen, 2006; Rosen, 2007). A simple search box, advanced searching, spelling corrections, suggestions for alternative searches, and faceted browsing are some of the features that the Primo solution offers end users. For a detailed description of the Primo search interface, see Sadeh (Sadeh, in press).

Figure 2. Primo search interface, including groupings for faceted browsing
How do I determine whether an item is relevant?
In addition to displaying metadata fields such as author, title, and subjects, the Primo system displays information such as an image of the book cover, the table of contents, and the abstract -typically originating from third-party sources - as well as information that is created by the community, such as ratings and reviews. Users can also look at the tags that other users have assigned to the item to see how the other users classify it.

![Figure 3. Full-record details with ratings, reviews, and tagging](image)

How do I obtain the materials?
Good discovery is, of course, mandatory for satisfying user’s expectations; however, it must be accompanied by an effective delivery mechanism. In Primo, each item on a result list comes with an availability statement that provides a link to the delivery service that is most appropriate. For example, for an article or a book that is available online, the system offers the user a direct link to the full text of the item; and if the item is an image, the system displays the image with the appropriate viewer. When an item is not available online, the system either displays the detailed holdings information, enabling the user to locate a print copy of the item, or offers a form with details already filled out, enabling the user to request the item through document delivery or interlibrary loan services. Besides the default service offered to the user on the basis of the item’s availability, the system displays other relevant services. For example, if the item is an article, the system displays a menu of services generated by the SFX link resolver, in addition to linking to the full text of that article.
How do I save results?
Users can save results for future use and can send results by e-mail. In addition, Primo integrates well with workspaces provided by non-library applications. For example, users can store records in other reference-management tools, such as Connotea, or simply bookmark items in their del.icio.us account.

How do I continue?
Often a research process only starts with an search; once users find items of interest, they follow various paths to discover related materials. Such paths can consist of citations, common subjects, and, in systems such as Primo, tags. By following tags, a user can find items that others consider relevant to a specific topic and therefore may be relevant to the user as well.

WHAT’S NEXT?
Putting a solution like Primo in place is only the first step in building the future library system. Next, libraries need to offer multiple entry points to such a system so that it will be accessible to users wherever they are and whenever they need it. Such entry points can be the institutional web site and course management systems, in which a search box would be embedded. A toolbar and a Google desktop gadget would provide users with a means to search without going to the Primo web site. Another course of action that libraries can take is to build on the Primo web services and extract information from Primo to use in other environments.

Libraries can also extend the scope of a system like Primo by harvesting more collections of more types. The harvesting of materials from institutional repositories of which users are not aware can enable libraries to expose their hidden collections and make them usable. In addition, libraries need to explore further expansion through collaboration: the architecture of Primo allows it to aggregate multiple Primo indexes. With this capability, institutions that implement Primo can optimize their efforts by collaborating to show their users one coherent search environment.

In her September 2006 lecture, Calhoun concluded with these words:

“The next generation of library catalogs will need to support the next generation of students, researchers, teachers and scholars at our universities and colleges. Satisfying their needs will require more modular,
interoperable library systems. Standards for connectivity and linking will be extremely important in making these
loosely coupled systems interoperate. Libraries and their information systems will find ways to not only co-exist
with the Googles of the world, but to take advantage of them to expose their rich collections more effectively and
to a broader audience.

We will need to define and carry out changes in how the local library catalog, and the books and journals it
describes, contribute to the next generation of information systems that support 21st century teaching, learning
and research. The local catalog, and the collection it describes, is not the sun. It is one planet among others that
support students and scholars.

Integration should be outward rather than inward, with libraries seeking to use and share their collections in
new ways. None of this will happen by libraries working alone. I believe we are about to enter an era of
unprecedented collaboration and partnerships among and outside libraries” (Calhoun, 2006b).

Today, at the end of 2007, we see the beginning of this process. More and more libraries have started by adopting
new solutions for addressing the rapidly changing needs and expectations of users. The next few years will show
which directions these systems take and how libraries will proceed in their constant quest to improve their
collections and services and to fulfill their function as trusted, knowledgeable sources of scholarly information.

REFERENCES

Antelman, K., E. Lynema and A.K. Pace: “Toward a 21st century library catalog”. Information Technology and

2007.

Calhoun, Karen: The Changing Nature of the Catalog and Its Integration with Other Discovery Tools.

meeting, 2006b. http://www.igelu.org/conferences/conference2006 stockholm2006/Session16/7-9-06-libraries-
catalogs-and-the-strange-new-digital-landscape-karen-calhoun.ppt/

Casey, Michael and LauraC. Savastinuk: “Library 2.0: service for the next-generation library”. Library Journal,

Coyle, Karen and Diane Hillmann: “Resource description and access (RDA); Cataloging Rules for the 20th


OCLC: Perceptions of Libraries and Information Resources: A Report to the OCLC Membership. Dublin, Ohio :

OCLC: College Students’ Perceptions of Libraries and Information Resources: A Report to the OCLC

Rosen, D.: Ex Libris PRIMO Usability Evaluations. Minneapolis, Minnesota : University of Minnesota Usability
Services, 2006.

Rosen, D.: Ex Libris PRIMO Round 2 Usability Evaluations Summary Report. Minneapolis, Minnesota :
University of Minnesota Usability Services, 2007.

Sadeh, Tamar: “Time for a change: new approaches for a new generation of library users”. New Library World,


WEB SITES REFERRED TO IN THE TEXT

Catalog of North Carolina State University. http://www.lib.ncsu.edu/catalog/
del.icio.us. http://del.icio.us
Encore. http://www.iii.com/encore
Facebook. http://www.facebook.com
Flickr. http://www.flickr.com
Innovative interfaces. http://www.iii.com
WorldCat Local. http://www.oclc.org/worldcatlocal/default.htm
YouTube. http://www.youtube.com

NOTES

1 The figures are based on an analysis of the statistical data published annually by the Association of Research Libraries (ARL), available at http://www.arl.org/stats/annualsurveys/arlstats/statsxls.shtml

2 The Next Generation Catalog Interest Group (http://www.lita.org/ala/lita/litamembership/litaigs/nextgencatalog/nextgencatalog.cfm) was formed by the Library & Information Technology Association (LITA) at the beginning of 2007. According to its web site, the group “gives
LITA a discussion space devoted to developments in the library catalog, its nature and scope, and its interfaces. It provides a forum for presentations and sharing of innovation in catalogs and discussion of future directions. Collaborations with other LITA interest groups, such as in the areas of emerging technologies and open source software, will provide opportunities for programming. See also the discussion by Coyle and Hillmann (Coyle & Hillmann, 2007).

3 The Primo groupings are based on the IFLA Functional Requirements for Bibliographic Records report (http://www.ifla.org/VII/s13/frbr/frbr.htm), with enhancements resulting from librarians’ feedback and input from development partners.