Subject Search Disconnect

Margaret Mellinger and Jane Nichols

Introduction

In 2005, OSU librarians considered creating online academic subject area portals, referred to as “Virtual College Libraries.” At the time, Virtual College Libraries were loosely defined as disciplinary information gateways tailored specifically for undergraduate use. Before building the Virtual College Libraries, a task force of librarians wanted to investigate the concept with undergraduate students through an online survey, focus groups and a usability study. As OSU librarians explored the complex issues of subject searching they encountered a long-standing disconnect between how libraries present academic subject related information and how students find it.

Many questions were generated throughout the needs assessment. For example, do users want simpler access to “best sources” for subject research? Should libraries filter information by discipline, subject area, or course? Would undergraduates use library portals that were tailored especially for them? How is subject searching defined and how should libraries address the subject search disconnect?

We define subject searching from the user perspective, that is, how users search for subject content or topical information, rather than how library tools organize information. Obviously, librarians have long enabled subject access to information. From the time libraries began arranging materials by subject in addition to title and author, new tools and techniques have been created to facilitate subject searching. A wide range of options have been employed: from Library of Congress Subject Headings and other controlled vocabularies; to pathfinders or other subject guides; to browsing collections arranged alphabetically or by discipline; to thesauri, taxonomies and hierarchical classifications; to online subject portals; and, most recently, to tagging, and graphical displays. This is an impressive array of accomplishments, but for at least the past twenty-five years, authors have noted the increasing failure of library subject arrangements to meet the needs of real users.

Librarians see that students are frustrated with libraries and have difficulty using them to access topical information. Students prefer commercial search tools to our catalogs, subject guides and library web pages. Catalogs are bypassed by “a large and growing number of students and scholars...in favor of other discovery tools.” Subject guides are largely unknown to students, even though they are now readily available on most aca-
ademic library web sites. Moreover, many library web pages offer a fragmented and confusing array of starting points for discovery. This model does not match users’ expectations for “one system or search to cover a wide information universe.” It is no wonder that a majority of undergraduates forego library search tools and start their research with a commercial search engine as has been reported by a number of studies.

Internet search engines shape users’ understanding, preferences and behavior; “the popularity of the Web appears to have influenced users’ mental models and thus their expectations and behavior when using a Web-based OPAC interface.” This observation is true for other online library search tools, as well. Rather than expecting users to conform to library models and methods, librarians must study users and track their preferences for subject searching, and for searching in general.

Methodology
OSU librarians administered an online survey, focus groups and a usability study to gauge OSU students’ searching preferences and skills. The survey of thirty-three questions assessed users’ familiarity with computer technology and their preferences for web page customization. The eighty-eight survey participants were evenly distributed among classes and the largest percentage of them represented the three largest colleges. Four focus groups, with twenty-four total students, tested the perception about the disconnection between the Libraries’ information presentation and users’ information seeking preferences. Ten people participated in a usability study. The usability study consisted of three questions and demonstrated users’ success when searching OSU Libraries’ web pages to find topical information and known items. The needs assessment results influenced our thinking about students and their search behavior.

Undergraduates’ Preferences for Subject Searching
Because of their extensive experience using the Internet, undergraduates are predisposed to the way Web search engines work. When undergraduates use library search tools, they expect to find a Google-like interface. The needs assessment helped us understand undergraduate preferences for starting points, search features and results lists and how those differ from current library practices.

Starting points
Users prefer a clear starting point; usually this means one search box that covers a wide universe of content. One survey respondent recommended that the library, “Make it more transparent when searching. I don’t care what external search engine it’s coming from. Just let me type it (a search query) into a search box.” Even though OSU Libraries’ web site included a single search box for a metasearch tool that searched several multidisciplinary databases, the needs assessment showed that not all users found or used it. Some found and used it but expressed frustration with the results and the search speed. The results were too specific and the search was too slow. The metasearch tool did not meet users’ expectations for a starting point or a single search box.

A long list of library search tools stymies users. At the time of the needs assessment, the OSU Libraries’ web site presented an “array of ‘silo’ systems” including the Libraries’ catalog, databases and e-journals. Students needed up-front knowledge to choose from the libraries’ list of tools. Many students in the focus groups and the survey expressed that they wanted the Libraries’ web pages to be easier to use. One comment, “sometimes I just get lost” neatly sums up their experiences. OSU is not alone; a recent MIT Libraries study found that “when confronted by the vast array of tools available to them, the students often exhibited uncertainty about what to do next.”

Undergraduate students prefer a “task-based information access system” which is “more useful compared to traditional approaches to information organisation and access.” Students experience confusion with library web sites that list search tools, because students do not recognize that these tools will help them accomplish their tasks. MIT librarians noted that undergraduates’ “tasks were overwhelmingly applied towards course-related goals.” They think books and articles—we give them catalogs and databases. They want content and we give them tools.

Familiarity is a driver for students when choosing search tools. Students in the focus groups reported that they tend to start their searches for topical information on the web. Using familiar search engines, students gain an overview of their subject and begin the process of placing their knowledge in context. If a student learns to use a particular online library search tool, such as Lexis-Nexis or J-STOR, and has success with it, the student has a tendency to use that tool as a starting place, even though it may not be appropriate for new information needs.

Search features
Undergraduates prefer to employ short natural lan-
language queries, without carefully constructing searches using Boolean operators or controlled vocabulary. In OSU’s usability study, all participants used natural language queries. Only one used a Boolean operator and none used controlled vocabulary. The test question may have encouraged this response because they were asked to search on the topic “endangered species,” a common phrase. However, research on online catalogs has demonstrated that “users build queries using only terms they have in mind and do not consult a thesaurus before a search is conducted.” Jansen and Pooh’s meta-analysis found little use of Boolean operators, and Amanda Spink notes that overall use of Boolean search is limited.

The needs assessment revealed that undergraduates sometimes experienced problems finding the right terms to use, and that they would welcome keyword suggestions and listings of popular searches. Large and Beheshti state that users encounter difficulties when “choosing suitable search terms to represent their subject interests.” Users tend to re-try their original search terms in successive searches rather than reformulate their queries. Spell check, “did you mean” and search term suggestions are highly useful for steering queries in the right direction. Students in the focus groups indicated that this type of immediate and actionable feedback was valuable. However, Spink reports that “more like this” relevance feedback was used only once in twenty queries.

Results lists

Users prefer acceptable matches to appear on the first page of results. Relevance-ranked results lists support this expectation. Only the top ten or so hits are viewed, and users seldom scroll to the second page of results. They prefer results with some indication that their search terms are present, such as the highlighting of search terms. This gives them immediate feedback of the success of their search terms. They also want to see brief summaries describing their hits. In opposition to these preferences, libraries have traditionally offered results in publication date or in alphabetical order. Because users view only one or two pages of results, a “combination of relevance and date ranking” should be considered. In library catalog and database results lists, it is not often clear to the user where their search terms appear, and summaries are not always available. Recently, some databases and catalogs have added faceted browsing, so that users can choose alternate ways to refine their search and see how their search terms are related to other possible terms. Also, “Library OPACs are looking to FRBR (Functional Requirements for Bibliographic Records) as an organizational concept that will help their catalogs intuitively present content delivered in multiple forms.”

Users want “good enough” sources, not necessarily the best sources available. The evaluation of “good enough” is based on a number of factors, the most important being ease of access and availability of full-text. The sources people choose to use follow the Principle of Least Effort; “people tend to choose perceived ease of access over quality of content in selecting an information source or channel...” Libraries have traditionally touted the high quality of their collections versus the open web, but quality alone will not bring users to libraries. Users will choose high quality library resources if they are easy to access in full-text online. Students will go beyond the principle of least effort if they perceive some value in doing so. For example, focus group comments showed that students were keen to know what their professors might recommend as worthwhile sources for their research. This finding may mean that the recommender systems currently used by Amazon and other websites would be favorably viewed in library sites.

We concluded that undergraduates’ expectations were not being met by OSU Libraries’ web pages, catalogs and search tools. Users do want easier access to sources for subject research. Would personalization and customization tools facilitate easier access and meet users’ expectations? Would students gravitate to “virtual college libraries?”

Undergraduates’ knowledge and use of customization tools

Personalization and customization tools seemed to promise a constructive approach to the “subject search disconnect.” We see personalization tools as the ability to tailor library resources according to a personal profile, and customization as the users’ ability to change the look and feel of a web page. Coupled with the view that personalization and customization could help solve the subject search disconnect was the general assumption that undergraduate students are technologically savvy. Current undergraduates are the Net generation, those “born with the chip.” Therefore, it seemed that they would be the optimal audience for personalized online library services.

There have been a number of library and institutional portal implementations. However, some of these
implementations, such as MyLibrary, initially had limited success. Other portal projects such as MyWelch seemed to succeed because of a very well-defined community of users. This prompted us to question the cost-effectiveness of developing such interfaces for large groups of undergraduates. Why weren’t these systems used as much as expected? To what extent would undergraduates use a subject portal? Although undergraduates are known to be technologically savvy, OSU librarians encounter many undergraduate students who are not as technologically adept as anticipated. We questioned the extent of OSU students’ technological skills and wondered what applications they commonly used.

Given these assumptions and questions, the task force surveyed OSU undergraduates to assess their familiarity with computer software and interaction with web pages, including the OSU Libraries’ web site. Survey participants were asked about their programming abilities; their use of Web 2.0 tools such as blogs, wikis, and RSS feeds; and their preferences for customization features on web sites such as MSN and Yahoo!, and in course management software packages such as Blackboard.

The general findings of the survey corroborated previous findings and assumptions and also delivered some surprises. It was not surprising that ninety percent of the students used Google as their primary search engine; that all of the respondents used word processing software; and that over three-quarters used presentation software. Students in our study heavily used communication tools such as e-mail, instant messaging and text messaging. Their technological skills and tools were deeply focused on communicating and the number of tools they used to do so. “Undergraduates are communicators” states the recent ECAR (Educause Center for Applied Research) Study of Undergraduate Students and Information Technology.

We were surprised to find that over three-quarters of the participants in the OSU assessment did not use blogs or RSS feeds. Also, a majority of students indicated that they were not familiar with the term “Wiki.” Apparently, the extent of students’ technological skills was less broad than we had expected, but we also see these results as a limitation of our study. User studies are a moving target. These results may reflect the timing of the needs assessment which took place in early 2005, before the upsurge of MySpace and Wikipedia.

As for whether students would choose to customize the OSU Libraries web site if given the chance, an underwhelming number of the survey respondents (twelve) replied that they would. Three students answered an open-ended version of the survey question by saying they thought it would be useful to choose what content appeared on their library pages, and two students thought it would be “cool” to change the colors. Focus group responses were also lukewarm to the idea of customization. These responses, in addition to the other information we had gathered led to our conclusions about the suitability of the Virtual College Library concept.

Suitability of Virtual College Library Concept
Our research concluded several things: a majority of students use non-library online resources, they find library tools difficult to use, and they are not especially keen to customize the Libraries’ web site. From the focus groups, we learned that undergraduates are very assignment-driven, focused on particular courses, and do not necessarily identify with disciplines or colleges. Students want to save time by quickly gathering the information they need for their assignments.

The task force concluded against pursuing the Virtual College Library concept as it did not seem to match undergraduate students’ preferences. Other determining factors in the decision were the students’ tepid response to customization and the high cost of purchasing, programming and maintaining portal software. As OSU Libraries’ technological infrastructure and as new web applications mature, we may revisit this decision, especially if we identify a suitable user community.

As an alternative to Virtual College Libraries, the task force chose to focus on creating course assignment web pages in collaboration with faculty. Course level pages more directly mirror students’ assignment-driven thinking. Other academic libraries have come to the same conclusion. Both University of Rochester and University of Minnesota Libraries have built tools to streamline the construction of library web pages for particular courses. The latest generation of MyLibrary offers “course-based access to library resources.”

Course assignment web pages are not the sole solution to the subject search disconnect. There is a broader disjuncture between students and libraries that must be addressed. Libraries will need to take a multi-faceted approach to crack this problem.

Recommendations
Users’ and librarians’ expectations for locating topical information have totally shifted. Alternatives to library tools are so effective that users can get what they need...
from them without learning to use the library. Moreover, librarians are no longer satisfied in serving only those who are willing to learn how to use our systems. Libraries now have a greater opportunity to make electronic journals, databases, datasets and other information more visible and easier to discover and access by a broader group of users, particularly those who would not have used the physical library. There are a number of approaches recommended by experts to bridge the gap between libraries and potential users, and to meet the expectations of both users and librarians.

Libraries can take several strategies to position themselves for the future. Each library will need to prioritize which to adopt based on their unique environment.

**Enhance library search tools functionality.** Clearly, libraries must match users’ expectations of searching for, identifying, and accessing relevant resources. Innovations are occurring in several areas. Some libraries have focused on the catalog, as Endeca at NCSU, PINES in Georgia, or Lamson Library’s WPOpac. Others have centered their efforts on creating a single search tool: Oregon State University’s LibraryFind is one example. The University of Rochester Libraries has concentrated on making articles easier to find and access. Many libraries are making their library web sites more user friendly, the University of Minnesota’s Virtual Library focuses on undergraduates.

**Make the library discoverable via the web.** Ensuring that your web page is in the top results of commercial search engines is one place to start. Abram recommends guiding users from web results back to your library’s content via open URL resolvers and OCLC’s library search. University of Pretoria Libraries describes their process of linking their openURL resolver, SFX, with GoogleScholar to simplify access to journal articles. Open WorldCat is making library collections “more accessible from the sites where many people start their search for information.”

**Enhance users’ experiences of library web sites.** There are many features that libraries can employ to enrich users’ experiences of library web pages. Options include the use of RSS feeds to highlight new books, and lists that display most popular searchers or top-requested books. Libraries should find points at which users experience confusion and provide context-sensitive help. Start creating online virtual library communities to mirror the social space of the library’s physical building. “Think about building a 3-D information landscape driven by the user.”

**Determine niche access to library resources.** As libraries web services mature, new ways of manipulating content will emerge. NCSU Libraries and University of Minnesota Libraries offer examples to emulate. Create subject groupings of resources to be searched by a federated search tool.

**Market library services to users.** Do not assume that our students know about us. They want us to market to them and they are accustomed to marketing directed at them. One survey comment encouraged: “More publicity! Let students know how to access the benefits of the library from remote locations.” Another student said she would have used the library web page had she known about it earlier in her career.

**Continue to partner with faculty** to integrate information literacy into the curriculum. Faculty can highlight library content by requiring it. Students respond to faculty-driven/course-driven pushes.

**Partner with others working on these issues.** Greater advances can be made through partnership. Some partnerships are already formed, such as RLG Programs Partners and the Digital Library Federation. Others could form out of smaller conferences. Access and code4lib are two examples. We can also leverage the work that is happening on an ad hoc basis across many libraries. If libraries could organize efficiently, and decide to work in consort, much could be accomplished.

**Culture Shift**

The culture of libraries is changing; but perhaps librarians can be more conscious of their own mindsets in order to move libraries and the field of librarianship forward faster. Changing the culture of libraries and the point of view of librarians is just as important as implementing the previous recommendations. Here are suggestions for creating a culture of change that will support libraries as we move forward.

**Create a culture of innovation.** Innovating isn’t just for the big players in the library field. Open source and social software invite us all to explore and participate. Libraries and librarians can and need to develop technical skills. In ten years, what aspects and how much of your library will be physical?

**Create more intuitive search interfaces.** Continue creating our own tools such as PINES, WPOpac and PennTags. Find ways to contextualize result lists with graphical interfaces or other means of placing results in broader and narrower context. Grokker and AquaBrowser offer examples of this multi-faceted browsing.
Create a participatory culture. Undergraduates are accustomed to participatory culture from their experiences with MySpace, YouTube, and multi-modal online games. In comparison, communicating via email and web forms seems so...one-sided. Creating venues for participation by our users will give us valuable first-hand information.

Create a culture of assessment. Are we meeting users’ needs? How do we know? Find ways to anticipate how users are changing and implement strategies to meet those changes. If each library worked towards meeting just one anticipated change, imagine what the cumulative effects will be.

Conclusion
The subject search disconnect is not a new phenomenon. But the convergence of users, librarians and tools will profoundly change our future, promising a new era for libraries. We look forward to seeing how libraries will integrate themselves into the broader information environment.

Notes
6. University of California, 2.
12. Yu and Young, 175.
17. Ibid., 303.
18. Yu and Young, 173.
36. This is in reaction to Derek Powazek, *Design for Community: The Art of Connecting Real People in Virtual Spaces*, (Indianapolis, New Riders Publishing, 2002).
43. PennTags description, http://tags.library.upenn.edu/.
46. Breeding, 9.