2008

Designing Search: Effective Search Interfaces for Academic Library Websites

Susan Teague Rector
seteague@gmail.com

Jimmy Ghaphery

Virginia Commonwealth University, jghapher@vcu.edu

Follow this and additional works at: http://scholarscompass.vcu.edu/libraries_pubs

Part of the Graphics and Human Computer Interfaces Commons, and the Library and Information Science Commons

Downloaded from
http://scholarscompass.vcu.edu/libraries_pubs/5

This Article is brought to you for free and open access by the VCU Libraries at VCU Scholars Compass. It has been accepted for inclusion in VCU Libraries Faculty and Staff Publications by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.
Designing Search: Effective Search Interfaces for Academic Library Websites


**Introduction**

Academic libraries support and provide access to an extensive array of information resources: the Online Public Access Catalog (OPAC), electronic journal subscriptions, online database search tools, digital library collections, and information repositories. In a 2004 report on the usability of metasearch, Elliot notes that “access to library resources…has become increasingly unmanageable for users and increasingly complex even for librarians.” While the mission of most academic libraries is to provide quick and easy access to information, they often face challenges in the actual design of online search interfaces to support this goal. Not only is much of the software in libraries proprietary and difficult to customize, but libraries must also weigh a number of design considerations before building search interfaces for today’s library users. These considerations range from understanding how users approach information in Web-based systems to having a good grounding in information architecture and usability concepts. A solid understanding of common user tasks as well as an appreciation for how search engines have influenced users’ information seeking and retrieval habits can greatly inform search interface design decisions.
There is little question that students today approach searching from a different standpoint than their pre-search engine predecessors. Faced with these inherent challenges in designing library search interfaces, how can libraries build effective search interfaces that support a wide variety of users and at the same time provide comprehensive access to assorted collections of electronic resources?

Virginia Commonwealth University (VCU) Libraries (the Libraries) began to research this question in 2007 within their campus community. Over a nine month period, the Libraries designed, implemented, and tracked the usage of two different search interfaces on their home page. Each interface allowed the user to choose between four different search targets. Search targets are defined as systems in which the Libraries’ users search for information (i.e., the library catalog, link resolver, metasearch tool and Web site.) In this case study, the authors discuss the complexities of designing search interfaces, outline findings from the 9 month study, evaluate the effectiveness of the two interface designs, and talk about future research in these areas.

Related Literature

The complexities of designing search interfaces are well-documented in library and information science literature and have been discussed since well before the advent of the World Wide Web or search engines. Scholars such as Bush\(^2\) identified issues in effective information retrieval as early as 1945 when he eloquently predicted that future researchers would have difficulty keeping up with the pace of knowledge creation. Bush wrote, “there may be millions of fine thoughts, and the account of the experience on
which they are based, all encased within stone walls of acceptable architectural form; but if the scholar can get at only one a week by diligent search, his syntheses are not likely to keep up with the current scene.”

Contemporary information scientists (Bates, Borgman, Kuhlthau, Marchionini) echo Bush’s thoughts on the intrinsic difficulties of finding information, and all advocate that understanding the user and his/her information seeking habits works to better inform information system design. In her 1996 study, “Why are Online Catalogs Still Hard to Use,” Borgman argues that “online catalogs continue to be difficult to use because their design does not incorporate sufficient understanding of searching behavior.” Borgman stresses that, when building systems, designers should understand and weigh both the “conceptual” and “mechanical” aspects of searching for “only when the conceptual aspects [of searching] are understood can the user exploit the system fully and effectively.” Borgman points out both in this study and in an earlier paper that, although these complexities can not be remedied by design alone, there are ways to alleviate user frustrations, such as simplifying user interfaces and building systems that “standardize access by providing a common interface to multiple systems…”

It is interesting (and somewhat discouraging) that these same discussions about standardized, uniform search interfaces are still commonplace in academic libraries more than 10 years after Borgman’s initial study. Discussions about searching often transcend the library literature and appear in books and articles on human-computer interaction, usability, and information architecture (Nielsen, Rosenfeld & Morville, Shneiderman
& Plaisant\textsuperscript{13}, Spool\textsuperscript{14}). Nielsen’s online Alert Box\textsuperscript{15} articles provide practical advice to Web designers. He often comments on ideal search interfaces for Web home pages, and suggests that the most effective interfaces are simple ones.\textsuperscript{16} Rather than offering complex search screens, Nielsen recommends that Web site designers consider \textit{progressive disclosure}, an older, yet effective application design device.\textsuperscript{17} Progressive disclosure dictates that users should be initially presented with only “a few of the most important options,” offering more advanced options to users upon request.\textsuperscript{18} This technique not only works to hide the complexity of search systems and thus, simplify interfaces, but also “improves three of usability’s five components: learnability, efficiency of use, and error rate.”\textsuperscript{19}

Like Nielsen, Miller\textsuperscript{20} also argues for simple search interfaces but within the library realm writing that “in the library world, we spend a remarkable amount of time and energy larding up our search interfaces with umpteen filters, Boolean pull-downs, radio buttons and so on.”\textsuperscript{21} After designing such elaborate interfaces, Miller remarks that librarians are often befuddled when users do not know how to use them, even after library instruction courses.\textsuperscript{22}

Currently, many academic libraries are migrating toward the simpler interfaces Miller describes and are exploring either single search box or federated search technologies for use on their home pages. At the time of writing this paper (2008), University of Oregon,\textsuperscript{23} Duke University,\textsuperscript{24} Oregon State\textsuperscript{25} and Penn Libraries,\textsuperscript{26} among others, had some flavor of single search box on their home pages. For users, these simple, uniform interfaces
could offer a welcome reprieve from the complexity of native search interfaces in proprietary systems. However, one should be cautious for unless the infrastructure of library federated search improves to provide better ranking, more uniform search results and tools for modifying searches, users may be lured back to the simplicity of search engines like Google. More than ever, libraries need systems that can not only present users with intuitive, simple to use interfaces, but can also support library patrons’ shifting mental models of online search systems.

Morville\textsuperscript{27} writes about the challenge of building online systems and supporting user mental models in his book \textit{Ambient Findability}. He describes the Web as a system that is constantly evolving, one that “allows our information seeking to grow more iterative and interactive with each innovation.”\textsuperscript{28} In order to support the various ways in which users now search for information, libraries will need to intimately understand online users and how they search for information and update systems to support these changes. Morville outlines future considerations for creating online systems writing:

Calvin Mooers reminds us that design of a useful information system requires deep understanding of users and their social context. We cannot assume people will \textit{want} our information, even if we know they \textit{need} our information. Behind most failed web sites, intranets, and interactive products lie misguided models of users and their information-seeking behaviors. Users are complex. Users are social. And so is information.\textsuperscript{29}
Designing effective search interfaces and uniformly presenting library collections on the Web can be challenging. Library Web managers must be fluent in a variety of topics that range from how humans cognitively interact with search systems to the basics of layout and design. As old modes for information discovery disappear and new ones emerge, it will be increasingly important for libraries to support these shifting mental models rather than holding onto complex search interface standards.

**Background**

VCU is a Carnegie Doctoral/Research University in the heart of downtown Richmond, Virginia. The Libraries provide access to a rapidly growing academic community of more than 48,000 students, faculty, and staff. The Libraries’ electronic collections are continually enhanced and the catalog holds close to two million bibliographic records. Its database and online journal subscriptions exceed thirty thousand titles, and as of 2008, the Libraries provide access to more than 175,000 electronic books. Other online collections include an emerging digital library with both audio and video.

When the Libraries embarked on a Web site redesign in late summer 2006, one of the key design goals was to build an effective, simple, and visible search user interface that supported a diverse community with differing research needs. The Libraries conducted numerous focus groups and design sessions with library users during the fall of 2006 in order to determine what type of user interface would be most useful. Librarians at the VCU Libraries ultimately decided that a search interface highlighting four of the most
heavily used search targets might give insight into how users would react to a single search box on the homepage.

The Design of Two Search Interfaces

As the Web team began to build the search interface, numerous design questions emerged, including the placement of the search box, which systems should be highlighted, whether advanced options should be offered, and labeling of the search box itself. After several iterations, the first search box was a simple search field for typing in queries and a drop down menu with various search targets. The search box was placed on the Libraries’ homepage in a left-side navigation bar (see Figure 1).

The following search targets were included in the drop down menu: ‘Catalog Search’, ‘Find Articles,’ ‘Journal Finder,’ and ‘Web site Search’ (see Figure 1). When the user entered a query for the default system, ‘Catalog search’ and clicked the search button, the system performed a keyword search through the Libraries’ OPAC (ExLibris Aleph). With
the ‘Find Articles’ option selected, the system queried the Libraries’ federated search software (ExLibris Metalib), performing a search on selected databases that included Google Scholar, InfoTrac OneFile, JSTOR, Science Direct and FACTIVA. A search submitted to the ‘Journal Finder’ queried the Libraries’ ExLibris SFX software used for accessing electronic journals. When the user selected the ‘Web site Search’ and entered a search term, the system used Google to search pages on the Libraries’ Web site domain (www.library.vcu.edu).

After the initial design of the search box was released, the Libraries created a script to monitor queries entered into each system from the search interface. Data about the location of the library user, whether the search was performed from the Libraries’ homepage or subsequent pages, the query the user entered, which search target the user chose (‘Catalog’, ‘Find Articles’, ‘Journal Finder’, or ‘Web site’) and a timestamp were captured and saved to a database table (see Table 1).

<table>
<thead>
<tr>
<th>Database Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>IP address of the user converted into the following description: ‘on campus’, ‘off campus’, ‘JBC staff’, ‘TML staff’,</td>
</tr>
<tr>
<td>Referrer</td>
<td>HTTP Referrer – location from which user began the search</td>
</tr>
<tr>
<td>Search Term</td>
<td>The search term the user entered into the search box field</td>
</tr>
<tr>
<td>Search Type</td>
<td>Which system the user decided to search: ‘Catalog’, ‘Articles’, ‘Journals’, ‘Website’</td>
</tr>
<tr>
<td>Timestamp</td>
<td>The date on which the search was submitted to each system</td>
</tr>
</tbody>
</table>

Table 1. Data fields captured from search box

From April 2007 to June 2007, the Web team monitored and reviewed search data, including the number of searches in each system as well as the most popular search
terms. After 3 months of examining logs and search terms, they determined the default search target for ‘Catalog search’ was used most often in searches; users were rarely switching between search targets in the dropdown menu.

In June 2007, the Libraries designed a second search interface, choosing a tabbed menu rather than a drop down menu to better expose the search targets. The text box for entering search terms was widened and the search feature was moved to the center of the Libraries’ homepage (see Figure 2). The new design was implemented on the public Web site before the start of the fall 2007 semester.

![FIGURE 2. Second quick search user interface design on library homepage](image-url)
Data Collection Methodology

Throughout the fall 2007 semester, the authors collected search data from the new design and compared it to the former design, using a script that captured user searches for each of the four search targets (see Table 1). To compare usage of the two designs, the authors focused on user searches from heavy research months for the Libraries, April and October 2007. For analysis purposes, the data from the two months was captured from the main database table and moved into a temporary table, where duplicate searches were removed. Duplicate searches were based on timestamp, search target (search type), search term, location, and referrer. (see Table 1). Only searches from the home page (where referrer is http://www.library.vcu.edu/) were analyzed.

The authors looked at frequency of searches in each system over the nine months, percentage of searches in each search target for April and October, and search terms (queries entered by the user). Search term data was organized into “tag clouds,” a popular Internet convention to highlight frequency of terms in an information space. Tag clouds helped identify how users were navigating each system as well as common search terms submitted to each system. The average length of search terms in each target was also analyzed and reported.

Results

After the interface design was changed from a drop down left-aligned search box to a central search box, the total number of searches increased by nearly 51% for the time period studied (see Figure 3). In comparing the volume and origin of the search terms
between April and October, it is apparent that the design changes made a significant impact on user behavior. In moving the search area to the center of the page and devoting more total space to it, the total number of searches increased dramatically almost 100 percent, with 17,593 searches from April compared to 35,037 searches from October. In comparing the two months, there are also differences in total usage of the homepage in the Web usage logs as reported by the Libraries’ Urchin software.\textsuperscript{30} For April there were 265,839 views of the homepage and in October this increased to 303,480. As this increase is considerably less than 100 percent, the increase in search activity cannot be attributed merely to increased site traffic.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{comparison_search_targets.png}
\caption{Comparison of Search Targets for April and October, 2007}
\end{figure}

FIGURE 3. Percentage of searches in each search target for April and October, 2007
It is interesting to consider the percentage of search to homepage views in the two samples. In April, search accounts for 8.00% of the homepage views and in October this number rises to twice that amount at 16.11%. This type of analysis could be very rewarding; however, it is difficult to ascertain what an ideal percentage might be. Certainly there are numerous institutional and technological factors, including the type of search technology available on the homepage and the variety of information needs the other parts of the homepage might serve.

The design changes also yielded the desired outcome of highlighting the variety of search options beyond the default “Catalog Search.” In moving to a tabbed design instead of the drop down menu, there was more screen space to indicate that a variety of search targets were available. Also, the tabbed design provided some space for small search hints and scope notes about each search option. Searches in the default Catalog option were reduced in the new design by almost 12%, from 72.44% of all searches in April to 60.52% of all searches in October. Conversely “Find Articles” was used almost 9.00% more in the new design (11.66% in April to 20.62% in October). “Journal Finder” usage also increased from 12.58% of all searches in April to 16.53% in October. The final option of “Web site Search” was relatively unchanged (3.31% in April and 2.33% in October). In short, the design changes do appear to highlight search options and give users some incentive to look beyond the default choice.

In focusing on the catalog search, it appears that users were significantly attracted to entering searches into the external search utility as opposed to clicking into the catalog.
where there are additional search options. In April, 16,716 users clicked into the catalog from a link on the Libraries’ homepage labeled “Catalog”, while at the same time 12,745 entered searches into the homepage search box for the catalog. For October, 13,219 users entered the catalog from the traditional “Catalog” link while 21,205 used the search utility as their choice. The desire for simplicity of search and level of comfort with traditional Web search engines discussed previously might account for this trend.

Another factor to be considered is that the external search utility has an unfair advantage in raw data comparison. For example, a user of the external search box may execute an unsuccessful search, and then click back to the library homepage to enter a second search, thus registering two searches in the Web logs. Regardless, the comparative numbers do indicate that a significant number of users were using the Catalog search option as their primary entrance into the OPAC.

A more central search box with a larger text area for queries also affected the length of search terms entered by the user. The largest increase in average search term length was in the articles search target in October, which increased from an average of 3.6 words to 4.1 words. These numbers suggest that the second design might be more appealing and easier to use than the first. As the authors felt that an in-depth analysis of the search terms themselves warranted further research, they did look closely at the longest search terms for each system in an attempt to support the hypothesis that users are able to distinguish between each of the 4 search targets. Clearly, the wider text area is more beneficial from a usability standpoint, but do users understand the systems in which they are searching?
An exploration of the top 1.00% of searches in each system (based on search term length) suggests that users understand the difference between an “Article” and a “Journal” search. However, what is not clear from examining this level of the data is whether users really understand what types of queries are appropriate for the catalog. In the “Catalog” search target there was a surprising number of full citation searches for either articles or books. Often in APA or MLA format, full citation searches were defined as those that contained author, title and year and could have been a citation for an article or book. The number of citations entered into the “Catalog” search target increased from 9.00% in April to 19.00% in October. This dramatic increase implies that users not only see the search box and believe it can parse a citation, but may think that the catalog functions like Google or other popular search engines. Taking this analysis further, the authors took the top citation searches from the Catalog for October and found that while a surprising number of citations return “hits” in Google (68.49%), only 34.24% actually returned the article on the first page of results for the citation entered. The fact that Google returned any results is an improvement over running the citation searches in the Libraries’ OPAC where almost all citation searches returned errors.

More in-depth analysis of the search terms could be very useful in building more robust library systems. The discovery of citation searches in the logs also issues a strong call for more advanced search parsing on the part of libraries. Users could clearly become frustrated by searching the Libraries’ resources and Google and finding disparate answers to their questions. As seen in the example, both cases more often than not, left the user without the article, book or journal they were seeking. More research is no doubt needed
to further examine the search terms; however, it is apparent from this level of analysis that the larger, more visible search box may lead users to falsely assume that the library search functions like an online search engine.

Discussion
Over the two months studied, the authors discovered that the second central search interface design on the Libraries’ home page proved to be more appealing to users than the initial left-aligned search design. Additionally, the tabbed approach used in the second design to highlight different search targets within the library domain made a significant difference over the drop-down menu in the initial design, as users were more enticed to choose search targets other than the default. The size of the search box where users enter queries was also important to the design of the interface. In the earlier design, the search box was too narrow, leading users to enter shorter queries than in the second design where the authors greatly expanded the search box size. While the authors discovered that the default search target was not selected as often in the second design, the catalog still remained the top choice for searchers in the library, making up 60.52% of searches in October. This begs the question of how often users really understand the difference between searching a library catalog, journal title, article or Web site. This also leads to a larger question of whether users really need to know the difference between these systems in order to find relevant information. A 2002 study by Cockrell and Jayne of how library users find articles on a library Web site, relates to how users distinguish between systems. They found that while the catalog “was the most frequent initial
choice...made from the library homepage...” users remained confused about “the role and content of the OPAC.”

Conclusion

As the amount of online information and the size of library collections continue to grow at unprecedented rates, it will be increasingly important for libraries to create simple and visible search interfaces for their users. This case study discusses what one academic library did to alter their homepage to provide a simple search interface to disparate collections of electronic resources. The authors hypothesized that a simple, central search box would entice users to search library collections and would help uniformly present library collections. Initial results indicate that the second central search box appealed more to library users. Despite the increase in usage, however, the question remains as to whether users actually found relevant and useful information.

Next Steps

This study helps the Libraries better understand their user population and its information seeking needs. Ideally, this type of research helps the Libraries migrate toward a single federated search box in the future. The authors of this case study found that a design incorporating a simple, central tabbed search box works well to uniformly present search targets within the libraries. Further research is needed, however, to ascertain whether the increase in usage of the central search box resulted in more effective searches and, subsequently, satisfied users. The authors intend to closely examine search terms, with an emphasis on studying variations in search syntax and relevancy of searches in the four
search targets. This case study coupled with a detailed analysis of search terms will likely illuminate the ideal search interface for the Libraries’ home page. For now, however, it is clear that the central search interface is an important feature for VCU Libraries’ Web site users.
Notes


3. Ibid., 105.


9. Ibid., 495.


18. Ibid.

19. Ibid.


21. Ibid.

22. Ibid.


28. Ibid. 60.

29. Ibid. 45.

30. For more information about Urchin Web Analytics Software, see http://www.google.com/analytics/urchin_software.html. As of 2007, the software was purchased by Google and is now under the Google Analytics umbrella.


32. Ibid.

**Bibliography**


Nielsen, Jakob. "Progressive Disclosure." (December 4, 2006),  

———. "Search: Visible and Simple." AlertBox (May 13, 2001),  


