by marshall breeding
Director for Innovative Technologies and Research
Vanderbilt University Libraries

The State of the Art in Library Discovery 2010

ALTHOUGH I SEE SOME WELCOME ADVANCES IN THE STATE OF THE ART IN THE REALM OF LIBRARY DISCOVERY INTERFACES, MUCH WORK REMAINS.

Resource discovery tops the charts as the foremost issue within the realm of library automation. As a new year commences, I see a more pressing need to accelerate the pace with which libraries deliver content and services in ways that our users will find compelling, relevant, and convenient. The evolution of the web advances relentlessly, and librarians have to work hard to ensure that our offerings stay reasonably competitive among the wares of the many information providers on the web. In this month's edition of The Systems Librarian, we'll review the current state of discovery tools for library content and consider some of the challenges that need to be addressed this year.

A Locus of Innovation

It's great to see significant progress in the development of new library interfaces. For the last 3 or 4 years, an incredible amount of creative energy has focused on crafting new tools to help library patrons navigate more easily through library collections that have grown to be more diverse and complex. Both commercial firms and open source initiatives have been busily involved in developing new offerings in this space. Commercial products include Encore from Innovative Interfaces, Primo from Ex Libris, AquaBrowser Library from Medialab Solutions, Visualizer from VTLS, Summon from Serials Solutions, EBSCO Discovery Service, Enterprise from SirsiDynix, and BiblioCommons. Major open source projects include VuFind, initially developed at Villanova University; Blacklight out of the University of Virginia; the eXtensible Catalog Project funded by the Andrew W. Mellon Foundation centered at the University of Rochester River Campus Libraries; and SOPAC, developed by John Blyberg at Darien Public Library in Connecticut. The surge of creative effort represented in these projects goes beyond what I have seen in the automation industry in any previous phase.

Not Merely Next-Gen Catalogs

Initially, these new tools were called next-generation library catalogs, but now I prefer to call them discovery interfaces. They aim to provide access to all aspects of library collections, not just those managed in the traditional library catalog, which is limited to the content managed by the integrated library system. It's all about helping users discover library content in all formats, regardless of whether it resides within the physical library or among its collections of electronic content, spanning both locally owned materials and those accessed remotely through subscriptions.

Discovery has been liberated from the confines of the automation systems used to manage library operations. One of the seminal breakthroughs in library automation involves

the separation of resource management from resource discovery. As long as each product that libraries use to manage content insists on providing its own end-user interface, we'll be plagued with a proliferation of user interfaces. In this model, which treats discovery as an independent activity, it becomes possible to channel the resources managed by multiple systems into a fewer number of interfaces, hopefully culminating in a single entry point to all library content.

Reconsidering the Role of the ILS

This approach repositions the role of the library's core integrated library system (ILS). The records managed by the ILS represent only one aspect of the content used to populate this new genre of discovery products. The separation from automation not only allows discovery products to address a more appropriate scope of content, but it also enables a more rapid, user-focused development strategy. Library management systems provide for the requirements of library personnel; discovery products serve library users.

For many libraries, we've moved past the time where a traditional online catalog of an ILS should be offered as the primary search tool for library content. Either the online catalog itself needs to be greatly modernized or the library should consider implementing one of the new discovery interfaces now readily available.

Modernizing the Look and Features

One of the key challenges involves delivering discovery interfaces that our users will want to use and that pass the muster of librarians naturally concerned with providing accurate, objective, and predictable search results. This new genre of discovery interfaces has revolutionized the library catalog, modernizing it into a form more consistent with other web destinations. Key qualities include a better visual design, relevancy-ranked results, facets for drill-down through search results, presentation of cover art, enhancement of records with summaries and reviews, and the ability for users to rate items or submit reviews. Some of the more recent innovations include incorporation of content supplied by users and other social data to enhance the discovery process and to provide resource recommendation features.

The context of the current state of the web demands that libraries shift from sites of complex navigation and fragmented information delivery to streamlined destinations crafted to guide users to high-quality resources. We are swimming in a very large pool of information providers. If libraries fail to offer more modern tools for discovery, our users will gravitate even more toward the commercial destinations. Libraries have a lot at stake in offering resources in ways that support our strategic missions, that enhance our branding, and that embrace the values of delivering content from objective sources rather than the sea of informal information on the broader web.

Equal Access to Content in All Forms

We must have discovery systems that do justice to the complexity of library collections. Today's library offers content spanning many formats and distribution models. For many libraries, books on the shelf represent a shrinking minority of library activity. Especially in the academic and special library realm, electronic content increasingly dominates our work. Even in public libraries, where print materials continue as the mainstay, electronic content and multimedia materials play growing roles.

Providing effective access to electronic content presents quite a challenge for libraries. For many years now, electronic content has steadily gained a larger place among our collections. The initial phase involved ejournals; ebooks seem to be taking off now. And it's moved way beyond text-based resources. Some of the most interesting content involves rich media in the form of streaming audio and video, photographs, and other images. We need to make it just as easy for library users to find materials among these electronic and multimedia resources as it is for them to find books.

The complexities of the tools that libraries use to manage content and the bundles in which we acquire resources should be transparent to our users. Acclimating users to the brand names of the publishers and providers that package scholarly and cultural content is not necessarily a vital component of information literacy. Rather, I think that the key goal of library discovery platforms lies more in offering streamlined access to the units of information held within—articles, long-form works such as monographs, as well as multimedia content.

Discovery, to be sure, is not the only concern. Libraries differentiate themselves through the provision of services surrounding the content they offer. We not only aim to supply access to information resources but also to provide research assistance, reading recommendations, and access to resources outside the confines of the local library's collection. As discovery products become effective in uncovering resources, the need for fulfillment services such as interlibrary loan, consortial borrowing, and document delivery increase. Libraries add tremendous value beyond what can be accomplished through even the best discovery platform.

Unify User Experiences

A typical library website presents a menu of many separate resources, each with its distinctive area of content and

user interface. I think that users might be better served through a more unified approach in the way that it delivers access to all these different areas of content. To the largest extent feasible, the library's web presence should offer users a seamless experience that presents a consistent interface, despite the use of multiple technology and content products behind the scenes.

One of the problems I notice involves the proliferation of search boxes on a library website. It's not unusual to see search boxes presented for the online catalog, for the website, on finding aids for articles and databases, on research guides such as implementation of the LibGuides service from Springshare that has seen booming popularity, and others. The scope of search, the way that results are returned, and the interface conventions vary tremendously among these search offerings. Multiple search boxes on the same site can confuse users. I have long advocated for integrating the content of the library website into the discovery product to avoid the need to offer a separate "search the website" service. The same principle applies to all of the other information management components that inhabit the library's web presence. Even though users may also need the features of discrete search environments for the later stages of research and analysis, I see great advantage to folding all this content into the library's strategic discovery platform.

Deeper Indexing

One of the most exciting developments on the discovery front involves emerging opportunities to offer discovery tools based on the full content of resources. The blending of the full text of journal articles and books alongside citation data, bibliographic, and authority records results in a powerful search experience. Rather than being provided a limited number of access points selected by catalogers, each word and phrase within the text becomes a possible point of retrieval. We've entered a fortunate time where the competitive and political obstacles that have previously prevented the ability for third parties to index full content have abated as technologies capable of handling massive indexes have been developed and refined.

Not that long ago, the idea of digitizing all of the books ever published seemed completely inconceivable. Today, that goal may not yet be realized but has come within sight. We've certainly reached a critical mass of full-text book content that can form the basis of powerful research tools. Google Book Search and HathiTrust count as early examples. The extent to which library discovery platforms can master full-text search and retrieval on a massive scale will impact our ability to compete with the commercial alternatives, such as Amazon.com and Google Book Search, that already leverage full-text search to great advantage. Even in cases where copyright restrictions prevent the display of digitized books and other materials, the digitized text lends tremendous power to discovery tools.

Serving Mobile Users

Libraries need to rapidly rally to deliver content and services to mobile users. The adoption of mobile devices in the general population of library users has reached a critical threshold. The increasing preference to access resources by small mobile devices ranks as one of those trends that libraries cannot afford to ignore.

The creation of apps for library resources has seen a flurry of activity in the last year. Many of the content and technology products used by libraries have begun to offer mobile versions. Through the use of tools such as Boopsie (www.boopsie.com), organizations can accelerate the creation of mobile versions of their websites, catalogs,

and other information resources. We read almost daily about the release of a mobile app for library-related services and resources. While it's great to see such an avalanche of library mobile apps, we also need to work toward a unified mobile experience for library users in the same vein as we've suggested for its web presence.

Aim for Web-Scale Discovery

It's now plausible to conceive discovery tools that address library content at the most expansive level of aggregation. Just like we expect general purpose search engines such as Google and Bing to attempt to comprehensively index all the content on the web, it's now plausible to construct library search platforms that include everything in our library collections. We use the term web-scale to characterize the discovery platforms that aim to manage access through a single index to all library content to the same extent that search engines address content on the web.

The current state of the art for library discovery lies in the realm of delivering products that embody this web-scale vision. We're seeing the emergence of services such as Summon from Serials Solutions, Primo Central from Ex Libris, the EBSCO Discovery Service, and WorldCat Local that aim to gather and index incredibly large bodies of library content. Others will inevitably emerge presently.

One of the most powerful capabilities of these services involves providing access to individual journal articles, a task that has long been relegated to external indexing services. The earlier phase of discovery product development focused on content managed locally within the ILS and other repositories. This round of development takes on the ambitious task of bringing in the vast realm of remote content accessed through subscriptions.

Web-scale discovery platforms will blur many of the restrictions and rules that we impose on library users. Rather than having to explain to a user that the library catalog lists books and journal titles but not journal articles, users can simply begin with the concept, author, or title of interest and straightaway begin seeing results across the many formats within the library's collection.

A great discovery interface should operate in a mostly self-explanatory way, allowing users to concentrate on selecting and evaluating the resources returned rather than struggling through the search tools that the library provides. Explaining the idiosyncrasies of the brand names of the publishers and providers from which we acquire information resources in wholesale often becomes the focus of information literacy and bibliographic instruction. Since so many library users consume the products we offer from outside our library buildings, having more intuitive tools to deliver library resources that do not require special training represents a valuable advance in the state of the art. The ability to assemble into a single index all the books, journal articles, and other collection components, in my mind, represents one of the most significant breakthroughs in library automation in recent decades.

Onward to Implementation

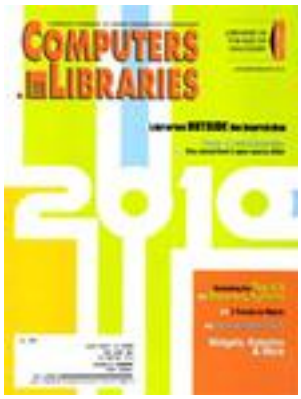
These strides in innovation on the resource discovery front will be made in vain unless libraries implement them expeditiously. I continue to observe sluggish adoption cycles despite the availability of many compelling products in both commercial and open source flavors. Even though quite a variety of new discovery products have been on the market for a few years, the vast majority of libraries continue to offer interfaces of an older, more traditional vintage. The costs involved with implementing new interfaces and the chronic paucity of resources perpetuate the problem of library interfaces lagging behind standard fare on the web. Yet I continue to feel a sense of urgency for libraries to offer the best possible tools for this critical aspect of their service delivery strategies.

Looking Ahead

Although I see some welcome advances in the state of the art in the realm of library discovery interfaces, much work remains. So far, only a minority of libraries have been able to implement one of these new-generation products. I hope to see progress in making these products more financially affordable and in lowering the thresholds in technical difficulty for implementation. We are also still in very early days of these new web-scale discovery platforms. While the concepts seem quite attractive, it will only be through the experience of library users that these products will either prove themselves or not. We should expect to see a continual process of filling in the gaps of content that products are able to address and ongoing refinements of the search and retrieval technologies involved. So while I'm excited about the progress made in the current cycle of development, it's paved the way for even more interesting innovations and the hard work of implementation, going forward. ■

Marshall Breeding is the director for innovative technologies and research for the Vanderbilt University Libraries, the executive director of the Vanderbilt Television News Archive, and the founder of Library Technology Guides (www.librarytechnology.org). His email address is marshall.breeding@vanderbilt.edu.

COPYRIGHT INFORMATION



Author: Breeding, Marshall

Title: The State of the Art in Library Discovery 2010

Source: Comput Libr 30 no1 Ja/F 2010 p. 31-4
ISSN: 1041-7915

Publisher: Information Today, Inc.
143 Old Marlton Pike, Medford, NJ 08055-8750