

The stage is set for a simpler search for users, but choosing a product is much more complex

THE NEXT

By Judy Luther
& Maureen C. Kelly

GENERATION OF DISCO

A casual Google search may well be good enough for a daily task. But if you are a college student conducting his or her first search for peer-reviewed content, or an established scholar taking up a new line of inquiry, then the stakes are a lot higher. The challenge for academic libraries, caught in the seismic shift from print to electronic resources, is to offer an experience that has the simplicity of Google—which users expect—while searching the library's rich digital and print collections—which users need. Increasingly, they are turning to a new generation of

search tools, called discovery, for help.

Libraries have been striving to respond to this challenge for years. The metasearch tools of the past decade—while different (and, ultimately, too slow)—were the first attempts to meet this user expectation by querying each of the databases a library subscribed to and returning a single set of results.

Enter discovery, which is modeled

LJ Explores the Big Tools

This is the third in a series of articles this spring devoted to new developments in major tools for libraries. The first, "Liverpool's Discovery" (*LJ* 2/15/11, p. 24), looked at a new search tool in action. We then tackled e-resource management with "Building a Better ERMS" (*LJ* 3/1/11, p. 22). Next up, read all about the many trends in ILS development in the April 1 issue of *LJ*.

on the Google-style approach of building and then searching a unified index of available resources, instead of searching each database individually. While Google's general index focuses on publicly available web content, these new discovery tools—including EBSCO Discovery Service and Serials Solutions' Summon, among others—provide unified indexes of the licensed scholarly publications combined with locally held content (like the catalog).

In effect, discovery tools make good on the promise of those earlier search solutions by shifting some of the IT management responsibilities to the cloud, streamlining search, and improving the relevance ranking of results. And users get to enter a single query—à la Google—to search the rich content of the collection with the speed they have come to expect. Still brand new, and in

Judy Luther is President and Maureen C. Kelly is a Consultant at Informed Strategies LLC, Ardmore, PA. The authors acknowledge the contributions of the librarians and vendors who participated in interviews conducted in the preparation of this article

action in only a handful of academic libraries, these tools are expected to transform search as we know it.

The challenge of findability

In a March 1994 *Wired* magazine article, nearly two years before the start of the research project that would become Google, futurist Paul Saffo made a prediction: "The future belongs to neither the conduit or content players, but those who control the filtering, searching, and sense-making tools we will rely on to navigate through the expanses of cyberspace."

That insight is as relevant to today's world of scholarly information as it continues to be for the other digital media that fill the Internet. When we look back at the early days of Internet search, we see that, while breadth of content was critical for search engines, the real success came from the algorithms that

by also searching library databases individually. Librarians hoped that metasearch engines, also referred to as federated search engines, would simplify searching across these databases, but they have fallen short of librarian expectations—owing in part to structural complexities, as connectors to each resource are hard to maintain. They have also disappointed users—due to their slowness of response, problems with relevance ranking, and inadequate handling of duplicates.

The new unified-index discovery tools offer great potential for simplifying scholarly search and making it more effective. As with all technology solutions, however, myriad details need to be sorted out in the move from concept to operational success. And the differences in how these tools are being implemented have implications for both libraries and for the publishers that supply the information.

SEARCH

- Simplicity of the interface.
- Quality of results, including relevance ranking.
- Ability to customize search and relevance settings.
- Availability of tools for navigating search results (such as clustering, facets, etc.).
- Ease of incorporation into existing institutional access tools.
- Support for new use environments, including mobile access and social-networking features.

FIT

- Ease of implementation.
- Compatibility with existing software and content.
- Responsiveness of the vendor and alignment of priorities regarding future developments.
- Overall customer support, including reputation and prior dealings with the vendor.

COST

- As a new service in addition to existing tools.
- Instead of other finding tools or delays to other upgrades.
- Justification in light of libraries' goals and objectives.

Content

To create a unified index, vendors need to secure permission from each publisher. While agreements take time, the amount of content included in the indexes is growing steadily. As more libraries implement discovery tools, primary publishers that have not yet agreed are likely to feel pressure from libraries that expect acquired content to be accessible through these tools.

While it's possible to determine which databases or individual titles are included, a detailed comparison of discovery services at the title level is an overwhelming task, as coverage of titles varies based on the depth of the archive and the currency of the content.

Local content in institutional repositories can be included with OAI (Open Archive Initiative) harvesting of metadata and ingestion of MARC records from the OPAC. Catalogers may weigh in on the fields from MARC records that are indexed, which can affect discovery. Librarians may also need to consider special collections or files with audio, video, and images.

VERY

brought the most relevant information to the top of the results list. When content is abundant, finding the right content becomes the challenge.

While we may settle for sufficient and convenient resources in our everyday lives, precision (just relevant documents) and recall (all relevant documents) are vital for scholarly information.

Because of its convenience, many people use Google as the starting point for information quests. But for scholars, it is only a starting point. Academic libraries are treasure troves of carefully selected collections and resources that Google does not consistently index. Google Scholar, a bibliographic index of scholarly literature, has attempted to fill the gap but has only achieved partial success, and its content coverage is unclear.

Until recently, scholars and students have been left to augment Google results

What to consider

Librarians have been examining these new discovery tools carefully, but comparisons have been frustrating because these products are new and enhancements are ongoing. Nonetheless, librarians have narrowed in on certain features and capabilities that are key to making decisions about these tools. Naturally, different institutions weigh each factor differently based on local needs and objectives, collections, users, and staffing. Leading factors are:

CONTENT

- Scope and depth of content being indexed.
- Richness (and consistency) of metadata included in the indexes.
- Frequency of content updates.
- Ease of incorporating local content, if desired.

THE BIG FOUR

LJ asked four vendors of discovery tools using unified indexes to highlight the key features of their products.



MINDY POZENEL

Director of OCLC WorldCat
Discovery Services

OCLC

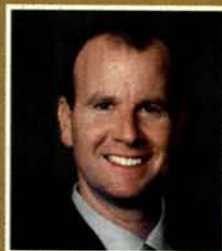
WorldCat Local brings the resources of any library together with the collections of the world's libraries. One search box gives users integrated access to electronic, digital, and

physical materials alongside intuitive delivery options for the materials they want.

WorldCat Local provides access to more than 740 million items, including articles from partners such as EBSCO, Elsevier, Gale, H.W. Wilson, and LexisNexis; the digital collections of groups like HathiTrust, OAIster, and Google Books; and the collective resources of libraries worldwide.

Users can find library holdings in search results from Google, Yahoo!, Bing, and Google Books using a portable search box that makes these library holdings visible on any website. They can also access materials from wherever they are with the WorldCat Local mobile interface.

WorldCat Local makes it possible for a library to consolidate discovery of many resource collections in a single place and expand its reach to connect users to resources beyond its collections to include the collections of OCLC member libraries around the world.



MICHAEL GERSCH

Senior Vice President
and General Manager

Serials Solutions

Summon pioneered single searchbox access across all library collections, providing a simple, easy, fast starting point for users. Librarians told us we

made their dreams come true. No surprise...we've all seen the statistics that show users want the quality of the library but prefer the simplicity of Google. We discovered through extensive user research that changing the perception of the library as "too complicated" takes a whole new technology, not a hybrid of federated search that patches the problems that drove users to Google in the first place. In a very short time, Summon has built a track record of delivering measurable value across the collections of more than 200 forward-thinking libraries. We have the usage analyses to back that up. We aim to bring users back to the library and keep them—so we created a technology architecture that accommodates rapid innovation. To wit, Summon's single, unified index is scalable: massive already—dwarfing any of the fledgling services—it will point to more and more content, regardless of its source, always eliminating information silos. It will never reduce its speed or accuracy in results. Hassle-free, flexible, and extensible, Summon gives the library impact by putting its most important asset—its collection—at the front door.

Librarians need to think about the role they want discovery to play in their libraries. If the tool is considered a place to start and a way to reach more library users, then complete coverage in a unified index may not be necessary for the undergraduate who is simply seeking "an answer." Others might decide it is even more important for the starting point to be as complete as possible. If the single search box is viewed as being the front door to all the libraries' resources, then librarians also need to consider how to present what is not included in the unified index.

Search

Discovery tools can effectively integrate libraries' resources with a single query across multiple databases that normally function as information silos. In essence, these tools create a unified space for integrating access to a diverse group of digital and print resources.

Vendors are fine-tuning their algorithms to adjust the relevance of search results. For example, many users expect searches to display a book itself before reviews of that book in a results set. With

journal articles, books, newspapers, conference proceedings, and other formats indexed together, vendors must consider how to weight the content types in response to queries.

The single search box lets users approach discovery tools with what they know and rely on filters to narrow the results. Scott Anderson at Millersville University, PA, which uses EBSCO Discovery Service, says, "We liked the idea that discovery tools reduce the cognitive load that the user has to know about the library." Facets that enable users to filter results by content type, subject terms, publication date, language, and other categories can also serve to acquaint users with the scope of literature available on a topic.

Librarians will need to decide where to place the search box on the library's website. Reference librarians may wish to identify additional resources not included in search results, or to highlight finding aids to help orient users. The single search box can be placed in course management applications such as Blackboard, or, ideally, into the student's workflow, wherever that may be. As Jennifer Duver-

may at Arizona State University (ASU), an early installation of Summon, notes, "We can't wait for the students to come to us; we have to go to them, embedding the search where they are working."

Fit

The desire to customize a system will vary by library. Librarians are inquiring about local control of system elements, from appearance (including labels for facets) to the ability to modify relevance ranking. Library staff with the technical capability to manage application programming interfaces (APIs) may want to add links to databases that are not included and establish profiles for disciplines. To incorporate local content seamlessly may involve including additional format types (such as government documents) or influencing metadata fields that the vendor indexes.

Holdings display and real-time data on availability of items will be important for those involved with consortial lending. While link resolvers appear to be compatible, data management may be required to address holdings or other factors that receive new visibility.



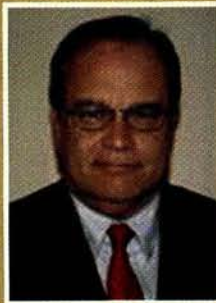
SAM BROOKS

Senior Vice President
EBSCO Publishing

All discovery services locally load metadata and full text from a number of primary publishers, creating a "unified index." Yet, these unified indexes do not include metadata from the overwhelming majority of secondary or full-text databases.

Other discovery service providers imply that the inclusion of these subject indexes and full-text databases is unnecessary, because "90 percent of the journals are covered" by the unified index. However, the inclusion of thin metadata is not remotely comparable, as secondary and full-text databases provide high-quality subject indexing and indispensable backfiles.

With the enormous results lists often provided by discovery services, it is critical that they deliver the best possible results on the first screen and provide strong data for facets to narrow giant results lists to more specific ones. The search experience is dramatically improved when comprehensive subject indexing from leading databases is available for searching and faceting. EBSCO Discovery Service (EDS) includes this crucial data via "Platform Blending"—the searching of leading subject indexes and full-text databases, when accessed via EBSCOhost, along with the unified index in an EDS search. EDS also offers "Optional Integrated Federation," which allows each end user to add-on databases that are only available via other platforms in an EDS search.



CARL GRANT

Chief Librarian
Ex Libris Group

The choice of over 750 institutions worldwide, Primo combines the breadth of scholarly content with a user focused interface and community-derived recommendations to support academic excellence.

Primo—together with its Primo Central Index of hundreds of millions of scholarly materials—offers a true one-stop shop for discovery and delivery, branded and customized to the individual institution's needs, with a choice of a local or cloud-based implementation. With its impartial content coverage, advanced relevance-ranking algorithms—configurable by the institution—and groupings of similar search results, Primo eases information overload, helping users focus on the most relevant materials that meet their needs. The built-in bX article recommender suggests additional relevant items by leveraging other researchers' selections of scholarly materials.

Primo provides needed OPAC functionality within the Primo interface such as requests, renewals, and fines; a variety of user services, such as a personal e-shelf; and access from mobile devices. The support of comprehensive consortia options, the ability to index materials from institutional repositories, databases, or other sources, multilingual support (including CJK), and a wide range of open interfaces enable libraries to integrate Primo into e-learning environments and social networks so the library can serve the end user wherever they are located.

Motivation and benefits

Discovery tools are evolving rapidly with input from partner libraries and through usability studies with end users. Early adopters were eager to address the growing number of databases and worked with their vendor partners to influence development. For libraries now looking at these tools, the sequence of planned enhancements may influence purchase decisions as a common feature set emerges to meet market expectations.

Discovery tools can leverage the institution's library investment through increased use of library resources, which can demonstrate value to provosts. Better access for undergraduates results in greater productivity for its users. Librarians at George Washington University, which chose EBSCO Discovery Service, have studied the options and see discovery services as "a tool that would reveal our content so that it's not hidden."

Many librarians who believe that discovery tools are a good first step in meeting the needs of the undergraduate student who is unfamiliar with the library's resources, discover it also meets the needs of researchers. Bryan Skib, as-

sociate university librarian for collections at the University of Michigan Library, which chose Summon, notes, "While very good for known-item searching, it's ideal for interdisciplinary research and for those who don't know what they're looking for or what databases to use."

Thanks to consortial agreements, small libraries can have access to as many databases as much larger libraries. As a result, the need for a single search across multiple databases is even more widespread than when federated search tools launched ten years ago. Schools with strong distance education programs acknowledge that today's learners value tools that support their operating in a 24/7 self-service environment.

Decisions and funding

While the decision to acquire a delivery tool is often led by a champion at the institution, the selection process tends to be shaped by environmental factors, such as synergy with existing systems or current sources of content and, of course, the available budget. A number of libraries that found federated search unsatisfactory have been eager to switch to or

acquire a unified index to offer users a better experience. The new discovery tools that incorporate federated search techniques use them as a complement to cover additional databases not included in the unified index and to provide a more comprehensive view of the libraries' resources.

Since buying a discovery tool is a new expense for libraries, funding has come from the systems or collections budgets, special allocations or possibly staff savings. At James Madison University (JMU), Harrisonburg, VA, which selected EBSCO Discovery Service, Jody Fagan says, "I can explain to library patrons why we don't have a particular database, but I can't explain why they have to use different search boxes for books and journals."

Some libraries form a team to conduct a thorough evaluation; it is likely to include staff from collection development, electronic resources, and bibliographic instruction, with input from cataloging. The team may also conduct side-by-side tests and consider factors such as ease of implementation, website changes, use of a link resolver, and customer support.

User opinions

Libraries are pleased with these tools and getting positive feedback from users. Early results show increased usage of library databases.

For experienced users and librarians at some schools, the reaction is mixed. "Some librarians love it and some hate it," observes Joseph Hafner at McGill University, Montreal, which uses OCLC's WorldCat Local.

According to library director Jonathan Miller at Rollins College, Winter Park, FL, which uses Summon, "We've had a business faculty who loves it and a philosophy faculty who hates it." Depending on the subject area, experienced users may prefer a familiar discipline-specific database that produces fewer search results, while those whose top-

ics cross disciplines are delighted to find new material.

Librarians have observed how they have gone from explaining the mechanics of search to focusing on evaluating search results. "Students are coming in with their problems rather than not knowing where to start," says ASU's Duvernay. "It's less about using the library and more about what they have found and how to effectively use it," observes Millersville's Anderson.

"We can start to move away from the mechanics of the database, and we can focus on the educational components and help students understand the difference in information objects," Anderson suggests. "Students need to think critically about what they have found rather than how to find it."

Future opportunities

As discovery tools become well established, there is the potential for them to help introduce users to new subject areas. For example, search results can be displayed visually, allowing users to navigate the literature and discover topics within it.

Vendors can incorporate new content types, such as e-textbooks and data sets, into the indexes and integrate them with learning systems. Video and image search will become smarter as more published content includes multimedia. Rapid adoption of mobile readers and tablet computers, such as the iPad, will help drive these developments.

Discovery tools are implementing mobile access, but resources that have not yet been "mobilized" can be unwieldy or unusable on a small screen.

All new technologies create both anticipated and surprising ripple effects. With the growing volume and diversity of electronic resources, it might be said that discovery tools are being introduced not a moment too soon. But what are the potential long-term implications of making scholarly content searchable through a unified point of access? Here are some key questions that librarians, publishers, and discovery-tool vendors will need to address.

THE LONG VIEW

LIBRARIES

LOST CONTENT? Will inclusion of a database in a discovery tool influence purchase decisions? Will publishers need to be included to survive?

FORMAT CONVERGENCE? Ebooks are joining e-journals online, along with multimedia. As we lose format distinctions in the digital realm, will a single interface become a logical approach to access?

DEMISE OF THE OPAC? Large libraries with more detailed MARC records and specialized collections may continue to view the OPAC as an essential tool, but will staff at smaller libraries that rely on MARC records from publishers and vendors question the value of the catalog?

WHAT COUNTS? With discovery and delivery more closely connected, will analysis of usage expand beyond PDF downloads of articles and incorporate more elements of web analytics, providing insights on user behavior?

PUBLISHERS

METADATA MATTERS? Early results indicate that while searches in the separate databases have declined, downloads of content from them have increased when discovery tools are used. Will publishers invest in enhancing metadata, given its importance for improving discoverability in prebuilt indexes?

ABOVE THE FOLD? As the volume of digital content expands, will publishers find themselves competing for relevance-ranking of their content on the first page of search results? Will search take its place alongside linking and social discovery of resources?

SEMANTIC ADVANTAGE? Are Abstracting and Indexing (A&I) services at risk in a world of full-text search? At this point, few A&I services provide metadata to discovery services. And certainly discovery services don't offer the same powerful indexing and sophisticated search interfaces that have been developed in these discipline-specific products. Librarians need to determine where discovery fits into their overall information literacy strategy and how significant it is for them to lead users to these more focused indexes. Is it important for students to discover these resources, as well as books, articles, and other content? Some discovery services do this well, others less so.

DISCOVERY

LONG TERM? Will the purportedly low switching costs for cloud-based services allow libraries to switch products as the services mature? Will selection of discovery tools be influenced primarily by content coverage factors? What will be the key points of differentiation for these tools as the market matures?

KEEPING UP? Patron-driven acquisitions are growing quickly in large academic libraries eager to offer readers more ebooks. Will vendors developing discovery tools keep pace with innovations in the consumer sector as discovery and delivery adapt to mobile readers?

BLIND SIDE? If Google addresses questions of content coverage in Google Scholar, or introduces other game-changing innovations, will Google Scholar become a serious competitor that dramatically alters the market?

CREATORS, NOT JUST FINDERS? Google Scholar and other search engines have become significant coproducers of academic knowledge by influencing search results and affecting what is easily discovered—with implications for usage, citation, and impact factor. Will publishers (and authors) recognize the potential impact of discovery tools on the community at large and address the risks as well as the opportunities?

OTHER OPTIONS

LJ's David Rapp on five alternatives to the unified index tools

AUTO-GRAPHICS' AGENT SEARCH Provides federated search that can be integrated into the XML-based AGent Web Services module or via a branded website. Staff can access statistics regarding database, website, and other usage. Users can also customize the look and feel of the interface.

INNOVATIVE INTERFACES' ENCORE Integrates federated search and harvested content. Functionality includes faceted search, and such social networking features as tag clouds and community reviews. Allows application development via its application programming interface (API).

SIRSIDYNIX'S ENTERPRISE Provides federated and harvested searching and faceted results. Social media and mobile support included. Users can save URLs of searches or RSS feeds to be informed of new library materials. Also allows library staff to create customized patron profiles.

TLC'S LS2 PAC Catalog module includes federated search with faceted results, as well as such social features as user-subscribed RSS result feeds, tags (which are fed back into the search index to help refine search results), and user reviews.

VTLS'S CHIVAS New discovery product slated for launch in May 2011, this is characterized as combining the functionality of VTLS's Visualizer discovery interface with that of its Chamo social OPAC to "produce a unified user experience, combining broad discovery across multiple resources with full OPAC access through faceted results."

As more publishers adapt content to personalized readers, the user experience will improve and could significantly change the way we work. Content providers will also increasingly use social networks to enable discovery through affiliation.

Going forward

While still fairly new, discovery tools are rapidly gaining content, adding en-

hancements, and growing their customer base. Libraries have adopted technologies that transform their services, and discovery tools are the next innovation. The unified index enables libraries to provide easier access to their resources at a time when mobile devices are beginning to change how we work. These discovery tools open the door for digital natives to encounter library-friendly services with a low barrier to entry.

When we look at the new discovery tools, we should remember what drives information decisions in our everyday lives. The new discovery tool providers would do well to note a recent observation by InfoWorld's Robert X. Cringely on why the iPhone and iPad have enjoyed such success: "The reason? Jobs' insistence on giving people what they really want: simple, intuitive products that are fun to use." ■

LIBRARYJOURNAL.

Webcast
Event

Want to know more about discovery and get your questions answered?

Join LJ for a special 90-minute webcast event with some of the creators, and users of the leading discovery tools. The webcast will further navigate the discussion of discovery as reported on in a series of spring 2011 LJ articles highlighting the development and implementation of these major tools. We'll look at how libraries are implementing discovery, how discovery is changing search, and flesh out the differences among the various products. We'll turn the conversation over to our audience, giving participants a chance to join in the discussion.

Save the Date!

Wednesday, May 4, 2011 1:00–2:30 PM EST

Be on the lookout for registration information online at www.LibraryJournal.com

COPYRIGHT INFORMATION



Author: Luther, Judy; Kelly, Maureen C.

Title: The next generation of discovery

Source: Libr J 136 no5 Mr 15 2011 p. 66-71

ISSN: 0363-0277

Publisher: Media Source, Inc.

160 Varick Street, New York, NY 10013

Library Journals LLC, a wholly owned subsidiary of Media Source, Inc. No redistribution permitted. To contact the publisher: <http://www.libraryjournal.com/>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden. The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.