Scholarly Papers and Articles   University Libraries

2013

Results of Web-Scale Discovery: Data, Discussions, and Decisions

Jeffrey D. Daniels
Grand Valley State University, danielsj@gvsu.edu

Laura Robinson
Serials Solutions

Susan Wishnetsky
Northwestern University's Feinberg School of Medicine

Follow this and additional works at: https://scholarworks.gvsu.edu/library_sp

Part of the Library and Information Science Commons

Routledge
Taylor & Francis Group

Results of Web-Scale Discovery: Data, Discussions, and Decisions

JEFFREY DANIELS and LAURA ROBINSON
Presenters

SUSAN WISHNETSKY
Recorder

A library with the best and most credible resources available is of little value if its patrons do not find the resources. Web-scale discovery products were created to provide a search mechanism that library patrons will actually use, and which will yield relevant results even for the most inexperienced of searchers. An overview of one such product, Summon, is presented by the company's product manager and by a librarian who helped to implement the product for his library.

KEYWORDS Google, searching, Serials Solutions, Summon, Web-scale discovery

At the end of the last decade, Web-scale discovery products began to surface in the wake of libraries' disappointing experiences with federated search engines. The new products, which combine all available resources into a single index for searching, promised to resolve the slowness and unreliability of the multiple searches performed in federated searching.

Grand Valley State University (GVSU), an early adopter of a Web-scale discovery system, is a comprehensive university with twenty-four thousand students on campuses throughout western Michigan. The library system, with a $4.5 million annual materials budget, about sixty-three staff members, over three hundred databases, over sixty thousand e-journals, and over six hundred thousand e-books, is a 2012 Association of College & Research Libraries (ACRL) Excellence in Academic Library Awards winner.

Presenter Jeffrey Daniels, who has been with GVSU since 1998, currently serves as Head of Technical Services and Electronic Resources Management. Daniels described the situation that led his library to become the first commercial implementer of Serials Solutions' Web-scale discovery product, Summon, in the fall of 2009. The library's website offered a number of search

options on its website, including its Innovative Interfaces Online Public Access Catalog (OPAC) and the Nautilus federated search product as well as various tabs for document delivery and course reserves. With a Web-scale discovery product, the library hoped to create a simpler, single-search-box search interface.

At the time, although EBSCO Information Services, Inc. (EBSCO), OCLC, and other companies had Web-scale discovery products in various stages of readiness, Summon was the only Web-scale discovery product available on the market. A major part of the implementation process for Summon, creating an up-to-date knowledgebase, was already completed, since the library was using Serials Solutions' electronic resource management (ERM) system. Summon was rolled out in less than a month, going live in August 2009.

Daniels recommended that any library considering—or already implementing—a Web-scale discovery system repeatedly ask itself five crucial questions, as his library did in the course of decision making about Summon. The following questions are important to address continually during the process:

1. Do we want Web-scale discovery?
2. Where and how do we want to present our Web-scale discovery?
3. Who is the audience for the Web-scale discovery product?
4. Should we teach it, and if so, how?
5. Is Web-scale discovery working for the library system?

DO WE WANT WEB-SCALE DISCOVERY?

It had long been clear at GVSU that its federated search engine was not working well for its patrons; in addition to maddening slowness, the system exhibited timeout and other connection problems, as well as sorting and relevancy problems. It was also clear that many patrons accustomed to a single Google search box were confused by the multiple search options on the library website. So the decision was made to pursue a product that promised to deliver the faster, simpler, comprehensive search that patrons seemed to want.

WHERE AND HOW DO WE WANT TO PRESENT OUR WEB-SCALE DISCOVERY?

The answer to this question was refined by the implementation team as time went on and evidence of patrons' use of the product was collected. Initially

it was felt that placing the Summon search box at the top of the home page would drive its patrons to use the product, which it largely did, but usability testing later determined that inexperienced users were still having a hard time figuring out where to start. On the redesigned home page, the Summon search box is larger and more central on the page. Other search options are still presented, but some options, such as "find articles" or "find books and media," are actually Summon searches with facets applied.

WHO IS THE AUDIENCE FOR THE WEB-SCALE DISCOVERY PRODUCT?

For the library at GVSU, the audience was patrons who were unfamiliar with the resources they would need for their research, mainly, first-year or second-year students, or anyone—even graduate students or faculty—doing research outside their discipline. Daniels praised Summon's database recommender and its ability to direct patrons to the best subject-specific resources for their needs.

SHOULD WE TEACH IT, AND IF SO, HOW?

Most of the classroom librarians felt that they should teach it. With less than an hour of class time devoted to freshmen, librarians felt it would be impossible to cover the intricacies of over 300 databases. Since Summon indexes database content, librarians find it more effective to steer new users toward Summon and describe some of the advanced searching techniques (which patrons can also use, later on, in subject-specific resources). The librarians also liked to demonstrate features that would impress patrons and get their attention, such as the ability to log in, save searches, and e-mail results; these "bells and whistles" helped to gain the students' attention and win their trust in the library as the best place to start their research.

IS WEB-SCALE DISCOVERY WORKING FOR THE LIBRARY SYSTEM?

It is certainly being used—sessions and searches in Summon have increased every year since it went live, topping one hundred thousand sessions and one million searches in the 2011 calendar year. The library also saw dramatic increases in use of its full-text journal collections and full-text aggregated databases; the use of abstract and index databases did decline, but not sufficiently to warrant any cancellations. The library had hoped to be able to recoup part of the cost of Summon by cancelling some resources, but the

usage data, for the most part, were as strong as or stronger than before the implementation of Summon.

Daniels directed the audience to a 2010 article by his colleague Doug Way, "The Impact of Web-Scale Discovery on the Use of a Library Collection," which gave more details on usage data at GVSU before and after the library's implementation of Summon.[1] He also advised librarians to look not only at Counting Online Usage of Networked Electronic Resources (COUNTER) statistics, or statistics provided by their Web-scale discovery system, but also at those from their link resolver, the mechanism that does the work to get users to the content.

Regarding the frequently expressed concern that catalog items would get lost in the index of a Web-scale discovery system and fail to appear in the results, Daniels referred to an upcoming article by his colleague Matthew Reidsma, the Web Librarian at GVSU, which found that while catalog items made up only 1% of all items in the Summon index, they accounted for 13% of search results.[2] Daniels mentioned his own upcoming article in *Technical Services Quarterly* about how the library's catalog was loaded into the Summon index.[3]

Daniels closed by saying that his library was continuing to refine Summon and its presentation on the library's website by asking the same questions, over and over, in pursuit of the ideal search experience.

Laura Robinson, Product Manager for Serials Solutions, began by describing her company. Created in 2000 by a university librarian, with librarians comprising nearly half of its staff, Serials Solutions shares the audience's passion for helping people find information, helping people make smarter decisions, helping researchers further their knowledge, and helping libraries evaluate their effectiveness in doing those things.

Polling the audience, Robinson found that few libraries have an employee specifically devoted to analysis or outcomes of the library; in many more cases, this task was just one of many jobs a librarian performed. Robinson cited an Association of Research Libraries finding that in 2007 and 2008 expenditures of a research library averaged about $26.5 million per library, and suggested that spending a small amount of that money on assessment and measurement of the library's impact might not be a bad idea. Robinson spoke of Carol Tenopir's 2009 study on the "value gap," in which patrons' perception of the value of the library has been declining even as library expenditures have been rising, and the 2011 University Leadership Council study *Redefining the Academic Library: Managing the Migration to Digital Information Services*.[4] She questioned whether traditional library metrics, even usage statistics, succeed in capturing the value of the library, and suggested that collection size—and even collection quality—is less important than ensuring access and connecting the patron to library collections.

Providing access is the mission of Web-scale discovery. In the course of research and development for the Summon product, Serials Solutions staff spent time with university students, and reached pretty much the same

conclusion as the *Redefining the Academic Library* report. It indicated that students believe the library has the most credible content, and that they prefer library resources for their academic research, however, they consistently begin their search with Google or other websites, because it is easier than searching the library's catalog or databases.[5]

Summon was created, said Robinson, to make searching library resources "feel like Google"; to index the library's catalog, electronic resources, and digital collections and "put it all into one search box" that would yield truly relevant linking results. Summon was designed to begin with a broad search, by default, since that is what students are accustomed to; the option to add filters is always there, once the searcher has begun.

At the University of Houston, which implemented Summon in 2010, a comparison of pre-implementation and post-implementation usage statistics showed similar results to those of GVSU: use of Elsevier resources went up about 75%, Institute of Electrical and Electronics Engineers (IEEE) resources nearly 50%, and Nature journals about 100%; the increase for Sage usage was the most striking, at about 225%.

But usage does not tell the whole story. We want to know how our libraries are impacting research, and impacting our students. Robinson suggested looking at several other assessment models recently discussed in library literature. Outcomes assessment involves looking at what an institution has to show for itself, such as the number of research publications produced, or how many times publications by an institution's researchers are cited. Contingent evaluation is calculating how much money is saved by library e-resource subscriptions or packages compared to obtaining the same materials with individual subscriptions or pay-per-view article purchases. And a similar model, return on investment, involves figuring out how much money an organization gets back, in the form of grants or tuitions, based on services and resources that their library provides.

An audience member asked Laura Robinson whether Serials Solutions regularly looked at the relevancy of various search engine results. Robinson responded that they do continually survey their own product's relevancy, as well as the relevancy of Google and other search engines, but they have no formal program to compare the relevancy of Summon results with those of other search engines.

Jeff Daniels was asked if he had found Summon rankings to be skewed based on the ownership of the resources in the index. Daniels said he had heard that concern before—that since Serials Solutions is a ProQuest company, ProQuest resources would be favored—but he had not found that to be the case.

However, he did find that certain resources, such as EBSCO databases, did not link properly in Summon. Since it is possible for librarians who administer Summon to choose their own rankings in overlapping instances of the same resources, the administrators at GVSU Library deliberately weighted ProQuest resources more heavily in these cases in order to allow patrons to

connect more easily to the full text content. Robinson responded to a follow-up question that there had been an ongoing dialog with EBSCO and other providers to obtain unique identifiers for linking. Daniels remarked that it is important for librarians, as customers, to keep the pressure on the companies to resolve problems.

An audience member questioned whether Summon really was faster than federated search tools; Daniels, after describing the single index that enabled Summon to perform one search, rather than multiple searches, assured the audience that Summon had no problems with speed.

An audience member at a library that already has a Web discovery tool said that one of the biggest challenges is getting librarians to accept the product. Jeff Daniels said he would remind them that the product is not intended for expert searchers or librarians, but for inexperienced students who do not know what a subject term is. Robinson added that it can be helpful to include all the librarians in the decision-making process right from the start, to get their input, answer their questions, and address their concerns.

Another audience member, in the process of implementing a Web-scale discovery product, spoke of librarians who wanted to heavily customize the product's relevancy rankings. Robinson expressed confidence in Summon's relevancy system and doubted whether extensive customization would be worth a library's time. Daniels agreed, saying that Summon's relevancy ranking was part of what a library paid for, and over time, the librarians at his institution had become comfortable that Summon worked well on its own without any intervention.

There were several questions about how link resolvers work with Summon. The answer was that it is up to the providers to provide direct linking to Summon, and more and more of them are doing so, using either the Digital Object Identifier (DOI) or a direct unique identifier. However, failure to link could arise from a problem with a link resolver, rather than a problem with Summon. The discussion led Daniels to remark that he hoped that Open URLs would disappear, as they so often result in linking problems.

Laura Robinson responded to a question on removing duplicates, explaining that Summon has a match-and-merge system that creates a single "super record." One audience member bemoaned the tendency of patrons—and even fellow librarians—to think of and describe the content in subscribed journals and databases as "free" to library users. Daniels remarked that there is probably no way for a library to fight that perception, although he routinely makes a point of showing his student workers the bill for the library's Elsevier package. What is more important, Daniels indicated, is that administrators at one's institution—the dean or president or provost—understand that resources cost money. Robinson added that students often recognize the value of library resources once they graduate and lose that access.

Robinson was asked whether students, after beginning with broad searches in Summon, actually use facets to narrow their searches; she responded that about 3% of them do so. Daniels added that the first three items in Summon results are selected over 80% of the time, and almost never will a user view a second or third page of results . . . just as is the case with Google searches.

An audience member commented that the high usage that the Summon product received at GVSU could be the result of the prominent placement of the search box on its website, making use of the product nearly unavoidable. Jeff Daniels responded that links to subject-specific databases are available as well, and by the time a student is a junior, he or she has learned about those databases. For example, students in the nursing program know what the Cumulative Index to Nursing and Allied Health Literature (CINAHL) is and go directly there. Summon is for those who need a place to start.

NOTES

1. Doug Way, "The Impact of Web-Scale Discovery on the Use of a Library Collection," *Serials Review* 36, no. 4 (December 2010): 214–220.

2. "Once Upon a Time: GVSU and Summon at Three Years," American Library Association Annual Conference 2012, Anaheim, CA, June 2012, https://speakerdeck.com/matthewreidsma/once-upon-a-time-gvsu-and-summon-at-3-years (accessed December 4, 2012).

3. Jeffrey Daniels and Patrick Roth, "Incorporating Millenium Catalog Records into Serials Solutions' Summon," *Technical Services Quarterly* 29, no. 3 (June 2012): 193–199.

4. Carol Tenopir, "The Value Gap," *Library Journal* 134, no. 12 (July 2009): 20; Carol Tenopir, "Measuring the Value of the Academic Library: Return on Investment and Other Value Measures," *Serials Librarian* 58, nos. 1–4 (2010): 39–48; and University Leadership Council, *Redefining the Academic Library: Managing the Migration to Digital Information Services* (Washington, DC: Advisory Board Company, 2011), http://www.eab.com/Research-and-Insights/Academic-Affairs-Forum/Studies/2011/Redefining-the-Academic-Library (accessed December 4, 2012).

5. University Leadership Council, *Redefining the Academic Library*.

CONTRIBUTOR NOTES

Jeffrey Daniels is Head of Technical Services and Electronic Resources Management for the library system of Grand Valley State University, Michigan.

Laura Robinson is Product Manager for Serials Solutions.

Susan Wishnetsky is Assistant Collections Librarian at the Galter Health Sciences Library, serving Northwestern University's Feinberg School of Medicine.