We are all aggregators (and publishers) now: how discovery tools empower libraries

Steven David Shapiro

With the proliferation of electronic resources such as bibliographic databases, eBooks, streaming audio and video, digital collections and online journals, the average library patron is often perplexed when it comes to navigating the thicket of information products available at larger academic and research libraries. What if libraries had the means to "aggregate" all of their content on one platform so that their patrons were spared the chore of sorting out what is most useful? In some sense, the technology to achieve this goal already exists. Discovery services like EBSCO's EDS, Serial Solution's Summon, and ExLibris Primo allow libraries to "aggregate" metadata from a large array of major publishers and vendors and make it searchable. While this does not constitute full-text aggregation in the strict sense of the term (although open URL resolvers, cross ref, and other forms of linking make full-text readily available), it does provide the illusion of being able to aggregate content like some of the commercial and non-profit vendors (e.g. EBSCO, ProQuest, JSTOR, Project Muse, etc.) (Shapiro, 2013). Yet, from the users' perspective, the question is academic. For all intensive purposes, we are all aggregators now.

Perhaps it was inevitable that libraries should become their own "aggregators" and publishers. New technologies (e.g. the internet, link resolvers, search engines, etc.) and market trends have empowered libraries (as well as individuals and smaller organizations) to compete with much larger corporate entities. Examples like Stanford University Libraries which founded Highwire (in 1995) and the Milton S. Eisenhower Library at Johns Hopkins University which teamed up with the JHU Press to create Project Muse (in 1995) are just two examples. Interestingly, Highwire and Project Muse recently announced an agreement allowing Project Muse to be hosted on the Highwire platform. There is logic to this evolving role of the library as content provider and aggregator. The serials crisis and the escalating price inflation of full-text content may very well accelerate these trends. Ironically, it is the vendor/aggregators (e.g. Ebsco and Proquest), themselves, who are providing libraries with the discovery service applications which could potentially liberate them from dependence on their all-encompassing full-text search products like EBSCO academic search complete and ProQuest central. If a library can collect all of its publisher content on one search platform and have full control over the quantity and nature of this material thereby meeting the needs of its institution in a more targeted manner then why subscribe to one of these "costly" full-text or niche databases. The counter argument is that discovery tools and the technology required for libraries to achieve this objective are quite expensive. The price tag for these technologies has declined tremendously over the past year or two. According to Robert Karen, Director of Procurement services for the Waldo Consortium, pricing for discovery tools has come down somewhere between 15 and 30 percent. More importantly – the services that enhance and support discovery services have come down in price even more due to the discounted packages offered by vendors that now bundle together the discovery tool, open URL resolver, A to Z list and other e-services (Robert Karen 2013, pers. comment, 13 June 2013). Despite these changes, however, libraries will likely continue to pay for "aggregated" content into the foreseeable future (out of convenience, vendor inducements and certain cost advantages). Nonetheless, these technologies will, in all probability, alter the relationship between database vendors and their library clients in ways that we cannot yet predict but will likely have significant implications.

At Montclair State University, we launched our EDS discovery service at the beginning of the spring 2013 semester. The implementation process took close to six months as we strived to improve the search experience and worked to refine the interface. The first step was completing the catalog, content, and customization and branding questionnaires. The catalog questionnaire required us to provide information on our current ILS system (product, number of bibliographic records, linking to catalog records, capturing field values, etc.) as well as the MARC fields and subfields we are using to identify some of the major e-resource formats in our catalog including e-books, audiobooks, streaming video, streaming audio and e-journals while the branding questionnaire gave us the opportunity to arrange the interface for our EDS search page. Unsurprisingly, the content questionnaire was the most time consuming of all the documents to complete and in many ways is the central element in constructing the core of the discovery tool. It is here that you select the resource content and metadata to populate your EDS profile. What is remarkable is the wide range of vendors that have partnered with EBSCO along with the breadth of bibliographic databases (including full-text databases and stripped down indexes), reference collections and streaming video/audio products that are represented. EBSCO claims to have developed more partnerships

than any of its competitors. While mainstay resources like Lexis-Nexis (without the news content), Science Direct, Project Muse, JSTOR, PsycInfo, Naxos Music Library and Encyclopedia Britannica are all part of the package there is also a healthy smattering of open-access resources capable of being integrated into EDS. This includes many intriguing and useful resources like OCLC's OAISTER which consists of over 25 million records from a variety of digital collections as well as the OAPEN Library (of digitized e-books), the Social Science Open Access Repository, arXiv, the University of Pittsburgh Repositories, DOAJ and Scielo. In other words, libraries now have the wherewithal to gather together, on one searchable platform, metadata from proprietary as well as open-access sources.

After creating our content profile and completing our other questionnaires, we began the process of uploading our catalog records into EDS. Most discovery services offer this option to libraries. We needed to do several test runs before successfully delivering the records intact. We identified a number of problems with the way our catalog records were rendered in EDS. For instance, items such as audio material, e-books, and e-journals often had the wrong icon assigned to them and multiple volumes in the same record often did not show up in the results list. Most of these problems we are corrected as they were picked up during the course of testing the search interface. Catalog uploads can be done as often as once a day using EDS.

Another important feature of EDS (and most other discovery tools) is the ability to incorporate records from institutional digital collections and make them searchable. This enables libraries to showcase content created at their home institutions (e.g. dissertations, digitized historical material, faculty publications, and journals) in the same way that many publishers (e.g. Sage, Elsevier, Springer, etc.) utilize portals to make their own content more widely available. An increasing number of academic libraries at campuses like Oregon State University (Summon) and University of North Florida (EDS) have also been using both discovery technology and institutional repositories to "publish" and "disseminate" local collections while making them easier to locate. This has the added advantage of more effectively publicizing the work of one's own faculty, staff, and students.

One of the major drawbacks of these types of services, at present, is the fact that the largest players in the industry, EBSCO and ProQuest, do not cooperate with one another. This makes it impossible for EDS clients to load metadata from ProQuest databases like ABI/INFORM and Historical NY Times into their profile (and visa-versa for Summon customers). For the time being, you need to create EHIS (EBSCOhost integrated search) connectors to manually add ProQuest and other incompatible databases to a search. Similarly, EBSCO and Ex Libris seem reluctant to exchange information according to an open letter posted on the Internet by the Orbis Cascade Alliance in Oregon (Orbis Cascade Alliance, 2013). Otherwise, the general trend seems to be going in the opposite direction with increased vendor partnering and collaboration. For example, Gale just signed an agreement with EBSCO in May 2013 to include metadata from its databases in EDS after a period of tension between the two companies abated. Another drawback, due to requirements of some major databases such as SciFinder and Factiva to have users register for their service, means that these important resources are not included in any discovery or federated search products.

Some of my colleagues claim that EDS has "overwhelmed" students with way too much information. A typical EDS search can turn up thousands of hits (which can be construed as a good thing). Yet, my colleagues would be hard pressed to explain why these students (and quite a few faculty) who have been experiencing "information overload" have been flocking to search engines like Google and Google Scholar while paying scant attention to the library. A recent study commissioned by Credo and Libraries Thriving of 1,502 students at 420 academic institutions that was released at the 2013 ACRL conference concluded that students most often (38 percent) consult search engines like Google when they begin researching followed by electronic resources (29.1 percent) and the library catalog (19.4 percent) (Schwartz, 2013). A similar study consisting of 2,500 students (undergraduate and graduate) appearing in Library Journal's Patron Profiles report for academic libraries also indicated that 73 percent students responded that their first instinct was to consult Google when they started a research project (Enis, 2012).

Discovery tools have the potential to reverse these trends. One of the first institutions to beta test a discovery service, Dartmouth College, started experimenting with Serial Solution's Summon in 2008. By the beginning of 2012, they found that Summon usage had doubled since 2010. All told, search activity had increased by a multiple of three or four times since Summon had debuted on-campus in mid-2009 (Rapp, 2012). A 2012 survey of 5,261 US academics published by Ithaka this year posed a question asking respondents to rate the importance of six leading library functions. One of the six functions they assessed stated the following: "the library serves as a starting point or gateway for locating information for my research". This indicator had been in a steady decline beginning with the 2003 survey and continuing through the 2009 report (reports are issued in three year intervals). Then something unexpected happened. The 2012 survey revealed a modest upswing in the same indicator (Housewright *et al.*, 2013). Faculty members were now more likely to perceive the library as the first destination for procuring information for research. Roger Schonfeld, one of the authors of the report, suggests that the advent and deployment of discovery tools may account for the subtle turnaround (Howard, 2013).

While there are no magic solutions for academic libraries wishing to make themselves more relevant, again, to their users in the face of multiple challenges (from the likes of Google and Google Scholar), it appears that the new generation of web-based technologies (i.e. discovery tools and institutional repositories) may provide the means for preserving the library's key role in supporting the research activity of colleges and universities. They also provide something more. The ability to enhance and redefine

academic libraries as "aggregators" and "publishers" of content tailored to the specific needs of faculty and students. All that is needed to achieve this is the will, vision, and patience to stay the course without succumbing to temporary setbacks (technical, organizational, fiscal, etc.). Though the technology has not fully matured, it has progressed significantly in a short time. In a recent article in Smart Libraries, Marshall Breeding indicated that he was "impressed with the ongoing improvements in library resource discovery services". In his opinion, there is certainly room for additional improvement in areas like identifying relevancy but the technology is definitely moving in the right direction (Breeding, 2013). What remains to be seen, however, is whether academic libraries and librarians will move quickly enough to take full advantage of the opportunities presented by these technologies before other unit's on-campus (e.g. the information technology department) take the lead.

REFERENCES

Breeding, M. (2013), "Smarter libraries through technology: discovery services improving", *Smart Libraries*, Vol. 33 No. 5, pp. 1-2.

Enis, M. (2012), "New LJ report closely examines what makes academic library patrons tic", *Library Journal Academic Newswire*, November 5, available at: http://lj.libraryjournal.com/2012/11/academic-libraries/new-lj-report-closely-examines-what-makes-academic-library-patrons-tick/ (accessed June 12, 2013).

Housewright, R., Schonfeld, R.C. and Wulfson, K. (2013), *Ithaka S+R US Faculty Survey 2012*, April 8, available at: www.sr.ithaka.org/research-publications/us-faculty-survey-2012

Howard, J. (2013), *Scholars Increasingly Use Online Resources but Value Traditional Formats Too*, The Chronicle of Higher Education, Washington, DC.

Orbis Cascade Alliance (2013), "Alliance board letter to EBSCO and ex libris", available at: www.orbiscascade.org/index/ebsco-and-ex-libris (accessed July 2, 2013).

Rapp, D. (2012), "Discovery at dartmouth", *Library Journal*, Vol. 137 No. 3, pp. 36-39.

Schwartz, M. (2013), "Credo survey: today's students struggle with TMI", *Library Journal Academic Newswire*, available at: http://lj.libraryjournal.com/2013/04/academic-libraries/credo-survey-todays-students-struggle-with-tmi/ (accessed June 12, 2013).

Shapiro, S. (2013), "Are aggregators an endangered species?", *Journal of Electronic Resources Librarianship*, Vol. 25 No. 1, pp. 77-79.

Steven David Shapiro (*shapiros@mail.montclair.edu*) is based at Harry A. Sprague Library, Montclair State University, Montclair, NJ, USA.