7-2020

The Silent Battle on the Budget: The Effect of Centralized Indexing on Collection Analysis in Primo and EBSCO

Casey D. Hoeve
University of Nebraska–Lincoln, achoeve@unl.edu

Christina Geuther
Kansas State University, cgeuther@ksu.edu

The Silent Battle on the Budget:
The Effect of Centralized Indexing
on Collection Analysis in
Primo and EBSCO

Casey D. Hoeve[1] & Christina Geuther[2]

1 University of Nebraska–Lincoln, Lincoln, NE, USA
2 Kansas State University, Manhattan, KS, USA

Corresponding author: Casey Hoeve achoeve@unl.edu University of Nebraska-Lincoln Libraries, University of Nebraska, 317 Love Library South, 1248 R St. Lincoln, NE 68588, USA.

Abstract

Significant cuts to the collections budget and a fire in the main library at Kansas State University transitioned the Libraries toward an increased focus upon electronic resources management. As electronic resource discovery continues to be hampered by market competition between Ex Libris and EBSCO library vendors, difficulties were identified in obtaining accurate usage statistics for resource renewal reviews, particularly EBSCO products. Kansas State University Libraries use the ALMA URM in combination with Primo discovery service, which experiences known resolver complications when paired with EBSCO resources. Technical services librarians and information technology staff tested several workarounds, but the question remains whether any proposed resolutions can be truly effective in avoiding collection usage biases and analytical discrepancies, resulting from exclusive indexing. Crunching the numbers, monetarily and analytically, the Ex Libris and EBSCO rivalry is an onerous battle on the budget and resource renewal process.

Keywords: collection analysis, discovery layer, EBSCOhost, Ex Libris Primo, indexing

Background on K-State Libraries' recent budget cuts and collection analysis

Kansas State University (K-State) is a mid-sized Carnegie-classified R1 university with campuses in Manhattan, Salina, and Olathe, as well as an online Global Campus. In May of 2018, a fire at Hale Library, the central library in Manhattan, KS, and a simultaneous serials budget crisis led to a deep analysis of collections for cuts and space planning ideas for a renovated building.

The State of Kansas and the higher education system among Kansas Regents Universities have been subjected to, and suffered several major financial setbacks in the recent decade. Myopic political decisions and mismanagement of taxation policies at the state level resulted in budget shortfalls, yielding reductions in higher education surpassing $100 million dollars (Rose 2018). Consequently, diminution of the University's budget subsequently affected K-State Libraries' budget, exacerbating previous complications of a continued flat library budget from 2011 onwards. Resource inflation by library vendors imposed additional financial pressure upon the Libraries' funding, resulting "in over $800,000 in decreased collections allocations, accounting for 55% of the Libraries cumulative budget [with] cuts of more than $285,000 annually to keep pace with inflation" (Hoeve 2019, 42).

The reverberating effects of state budget reductions inevitably converged to a crisis point in the Content Development Department, as the cost of information resources increased above the monetary threshold to maintain a balanced budget. Therefore, the Libraries were compelled to implement a cancelation project to un-encumber money, and focus upon retaining core databases and journals to support research and the curriculum. A moratorium was also placed on monographic purchases, except for selective cases in the areas of course reserves, non-interlibrary loanable research material requests, and seminal works for program support and accreditation.

Lacking an associate dean and a department head overseeing collections, the Content Development Librarian for Arts, Architecture, and the Humanities devised a review system that relied upon a combined analysis of quantitative usage statistics and qualitative feedback

(Hoeve 2019). Usage data was downloaded from vendor portals and combined into spreadsheets, documenting use, cost-per-use, and content coverage, providing a readable document to serve as a decision-making tool for breaking apart large journal packages, and selecting critical journals for supported areas of study.

In the following year, electronic databases, electronic journals, and print journal subscriptions were evaluated for renewal or cancelation. After having downloaded usage statistics covering the past three years from vendor portals the prior year, content development librarians noted that in many circumstances data from prior years was inconsistent. This was especially the case for the evaluation of our EBSCO databases.

Discrepancies in the data caused librarians to question the accuracy of statistical gathering by vendors and library systems. The discovery, and consequently, the analysis of patron use of K-State Libraries' online collections is dependent upon the Primo Central Registry, which does not facilitate discovery of all products, most notably those on the EBSCOhost platform (Orbis Cascade Alliance 2020). Therefore, any analysis would have an innate bias to collection development, and Librarians must mediate their collections through competing tools, including multiple workarounds from third parties to make strategic renewal and cancelation decisions.

According to Breeding (2018) 48% of R1 institutions use Primo and 37% use EBSCO Discovery Service. Add the 19% base from the same study that identified as Proquest Summon customers, and there is now a definitive dominance in the system market by the Proquest Ex Libris acquisition. Thus, incompatibility with the EBSCOhost product suite is a growing problem for libraries.

Implementation of ALMA URM and Primo library discovery service

K-State Libraries implemented the Ex Libris Alma unified resource management system in 2015, prior to the ProQuest merger with Ex Libris, and so Ex Libris Primo was the sole compatible discovery layer option available at the time. Over the years, this instance of the Primo product has been heavily customized by library information

technology staff. As Primo updates, more customizations are made or innovated as the library learns what electronic resources need work-arounds. Over time, products on the EBSCOhost platform gained particular attention among the staff, such as discovering the failed attempt of the Orbis Cascade API to yield compatibility between EBSCO indexing and Primo. Given this information, it was determined that this API would not be usable for the backend of K-State's locally hosted Primo.

The "Jane Hale" (a generic persona) electronic resource access ticketing team within technical services petitioned for databases with regular discovery problems to be included on an upper-left section of external links that would direct the user from Primo to those resources directly. Another attempt was to designate the administrative tool for EBSCOhost databases to conduct a multi-database search from a database directory A-Z list. The most reliable resolution seems to be educating the library staff and patrons that no matter the workaround, there will be problems with EBSCOhost product discovery.

Peculiarities with EBSCOhost and various workarounds for Patron-Level discovery

With the acquisition of Ex Libris and its products by Proquest, the Ex Libris portion of the company handles workflow and discovery for the library systems it sells (Breeding 2018, 13). Created in 2006, Primo is the flagship discovery system, currently utilized at K-State and soon to be replaced by the Central Discovery Index. Primo discovery at K-State reflects three sources: from the unified resource management system ALMA, our institutional repository K-REx, and Primo Central. There is very little autonomy for a librarian with a Primo Central arrangement.

To activate an item in the Primo Central Registry, the librarian searches by provider (interface vendor) and selects "on" or "off" for listed resources. Whether the resources are listed depends on the relationship between Ex Libris and other vendors. Entries to the Ex Libris Primo KnowledgeBase cannot be modified without creating a ticket for Ex Libris personnel to make the adjustments. Ex Libris regularly reports what additions have been made as options to

activate, but until then K-State Libraries continues to depend on patrons using a database A-Z list maintained through SpringShare LibGuides. EBSCOhost is not a vendor that collaborates on metadata for the Primo KnowledgeBase, as a continued disagreement exists between the vendors pertaining to where file transfer of the metadata will take place.

The institution leading discussions toward an effective resolution is the Orbis Cascade Alliance, which has documented its communication between the rival vendors since 2013. An API for an EBSCO Search Plug-in was created after greater awareness of this problem. Although Ex Libris provides instructions to implement it on their website, K-State Libraries, as well as the SUNY Libraries Consortium (SUNY 2019) found the solution does not work without error (Ex Libris Group 2020). In fact, after testing, the staff decided overwhelmingly that putting the functionality in place would be a misnomer. Much of the mapping between EBSCO Type to Primo Genre is unknown or inaccurate, and that is evident on the very instructions for implementation. The Consortium of Academic and Research Libraries in Illinois (CARLI) also supports these claims, identifying that the EBSCO API is problematic for searching instances in Primo, citing issues of inconsistent search results based upon user location and incompatibilities with Primo search fields (CARLI 2019).

In spring 2018, years after K-State's unsuccessful test of the EBSCO API solution to Primo, the team responding to tickets generated for electronic resource access issues decided to settle for a canned response: The tickets for EBSCO database problems when searching the Primo discovery layer would simply state that there would be no resolution anticipated between EBSCO and Ex Libris. In fact, the workflow documentation for electronic resource access resolution was modified with a separate offshoot for problems related to EBSCOhost products after routing through Primo. The reaction of staff external to the department was as expected: it was as if librarians were giving up on the patrons. In spite of several efforts to educate the public services departments on this issue, tickets continued to pour into technical services asking for tickets to be submitted to the competing vendors. In the three-year period from 2016 to 2019, technical services received 74 separate tickets related to ongoing metadata issues between EBSCOhost databases and the Ex Libris Primo discovery layer.

In 2019, the K-State team assigned to oversee local Primo discovery introduced a box of external links to the Primo search results page. At the same time, Google Scholar was also added to this link box, but with a different functionality. The search input within Primo populated into the Google Scholar search when clicking on this option; however, due to EBSCO's limited metadata functionality with Primo, the link would lead to a blank page providing various EBSCO databases to select from to continue a search. Essentially, the effect of the new EBSCO search button as an external link simply referred patrons away from attempting a search which would have, if any, only a mixed success of results.

EBSCOhost products continue to appear in a Primo search to this day, but often those links to full text are broken or lead to incorrect results. Library User Services requested that the external link lead to EBSCO Multiple database search in order to simplify the search process, however federated searching is not as effective. In federated searching, the common denominators of the various indices are used, and thus are ineffective in comparison to using a database as the vendor created it, especially for specialized knowledge. It is not to the extent of the relevance expected in a normal database search. Instead, the results are not relevant to make a COUNTER record or result clicks, and the record views are not going into statistics counts.

According to a study of library discovery systems, "Though important, the broad search service powered by an index-based discovery service may not be the primary way that most users gain access to library resources. Multiple research studies reveal that only a portion of research begins at the library website" (Breeding 2018, 28). There is considerable value for a patron in a specialized discipline to begin a search within their designated library resources including subscription databases. Searching within a smaller quantity of indices corresponds to broader application of jargon or otherwise unique metadata in the index. Additionally, the errors in search results are minimal and attributed to the vendor side of resource provision. K-State provides patrons with this more individualized search method by a LibGuides A-Z Database Directory that refers to applicable disciplines and a short description of the resource. Furthermore, this workaround is the most easily customizable for patron level discovery.

Exclusivity of content & the problem with statistics

While vendor exclusive content continues to remain a common practice, inconveniently serving as an obstacle to equitable information access (Conrad 2018), the problem at-large exists when exclusive content is coupled with exclusive discovery service indexing. The Ex Libris-EBSCO impasse is particularly noteworthy, given the large market share held by both providers (Breeding 2018, 28).

EBSCO's historical unwillingness to share indexing records (Krumenaker 2001) with other discovery services led to a salient outcry in 2018–upon EBSCO's agreement with the MLA International Bibliography to become the sole host of the database. Previously hosted through several other vendors (including ProQuest), EBSCO made the conscious decision to remove indexing records from all other discovery services except EDS (Horava 2019, 15).

The incompatibility of infrastructure increases entropy within a shared information environment (Hoeve 2018, 105), leading to poor functionality and unpredictable access (or loss of access) to the user. It was noted that the lack of discovery caused many libraries to discontinue their subscriptions to MLA (Horava 2019; Conrad 2018), as the EBSCO API was not deemed a sufficient solution to retrieve relevant content. Given that many EBSCO products are considered crucial information sources, exclusive indexing has the potential to adversely affect research and university programs.

These claims are corroborated in a review of three web-scale discovery services with equal access to Pubmed index data. It was demonstrated that there was not enough difference between Primo, Summon, and EBSCO Discovery Services to determine the best discovery service (Hanneke and O'Brien 2016, 113–114) (Stovold 2017), suggesting the exclusivity of indexing records causes the "aggregation [of titles] to become fractured" (Krumenaker 2001), negatively affecting access to databases. Ex Libris and EBSCO have provided conformance statements for the National Information Standards Organization Open Discovery Initiative (National Information Standards Organization 2014), in which "the content provider makes available to Discovery Service Providers core metadata and underlying full-text/orginal content for complete offerings" (EBSCO 2019, n.p.). However, previous complaints by Ex Libris and the Orbis Cascade Alliance regarding EBSCO's practice

of exclusivity have not yet been adequately resolved (Thompson 2014, 132–133; Horava 2019, 17). Given the fact that users have little or no choice but to use the ILS and Discovery Services selected by a library (Deodato 2015, 20), exclusive indexing disadvantages patrons and impedes the statistical analysis of database usage.

Regarding analytics, multiple platform hosting requires extensive and laborious analytics work to evaluate usage. Statistics must be gathered from all vendors that host the content, and report merging is necessary to combine and analyze usage to make informed decisions. Such was the case for K-State Libraries, which had access to MLA on both ProQuest and EBSCO platforms for approximately six months. Librarians concluded that both platforms should be made available to users, as the transition from one platform to another was necessary for proper information literacy instruction.

Furthermore, changes in ownership cause additional difficulties for libraries when determining ownership of content. Journals or databases that change providers (Branch, Hovenden, and Upton 2019, 112) may receive updates through the ALMA Community Zone and EBSCONET; however, if indexing records are exclusive, the records can become inaccurate in link resolvers, as notes identifying perpetual access may not transfer over, or records may disappear from the catalog entirely (Kennedy 2017). Librarians must remain vigilant, and spend countless hours of labor to ensure records and license records are correct. At K-State, the Electronic Resources Librarian was tasked with documenting perpetual access content, a process that necessitated two years to complete given competing work priorities.

Considering how frequently libraries have brought up this issue, organizations, publishers and vendors should be well aware that exclusive agreements cause complications and not only change the dynamics in how patrons use platforms but also influence the relevance of content retrieved. When content becomes inaccessible, the research and its potential are diminished. Egregious squabbles for profit and market share cause significant consequences that librarians must carefully navigate to curate the best and most needed collections. This becomes an arduous endeavor, as instruction, and qualitative and quantitative statistics must be skillfully used to surmount these obstacles.

Workarounds for Alma Analytics and other collections analysis methods

In a collection development environment with shrinking budgets, the prevention of index record sharing can result in detrimental effects upon disciplines that rely upon specific vendor databases. With fewer opportunities to be discovered, usage statistics are presumably lower than they should be, as access to content is not equal through Primo Central. Such complications of low use can generate issues of higher-cost-per-use, thus targeting necessary databases for cancelation. It is then incumbent upon librarians to make a case based upon qualitative evidence, faculty outreach, and flawed data-driven arguments.

K-State currently utilizes ALMA Analytics to evaluate print use of monographs and journals. As vendors become COUNTER compliant with SUSHI harvesting capabilities (COUNTER 2019), accounts are established to automatically transfer usage statistics into the system via API. However, during journal cancelation projects, it was discovered that ALMA Analytics reports derived from SUSHI usage data were different from usage data generated from vendor administration accounts. The discrepancy between SUSHI data and vendor portal data was confirmed in an email chain on the SERIALIST LISTSERV, in which, the COUNTER Executive Board explained that the COUNTER Code of Practice allows for a reliability window between 3% over-reporting and 8% underreporting (Emery 2020).

The EBSCO API does not provide optimal functionality with searching in Primo, so it was further questioned as to how reliable the usage reports for EBSCO resources were compared to Primo Central indexed databases. As the Libraries rely heavily on usage statistics to make renewal or cancelation decisions, multiple methods had to be employed to identify outliers for additional scrutiny (Lamothe 2014), or analyze general usage trends for non-indexed databases.

To work around non-Primo Central indexed databases, other forms of analysis were used to gain insights into usage trends. While not completely accurate, usage data (either COUNTER or non-COUNTER) reports were pulled from vendor administration portals spanning a three year period to examine average use, looking for trends or variances in the data. In addition, the Libraries A-Z Database List and LibGuides are hosted on Springshare, allowing content development

librarians to view the number of clicks on a resource link every year. Interlibrary loan reports were also requested to determine how many in-house loans were fulfilled from local collections in non-indexed resources (Musser and Coopey 2016, 646). As a last resort, database renewal priority was based upon potential need for program accreditation, and qualitative feedback from Libraries and university faculty on the importance of a resource, and how often it is used for research, course instruction, or information literacy sessions.

The practice of exclusivity regarding indexing causes collection development analysis of usage statistics to revert to making judgments with flawed data or alternative, less informative, and time-consuming analysis through outdated methods. This places academic departments that rely on these resources in a precarious position, with non-indexed resources requiring greater justification each year to prevent cancelation. Based upon literature and case study, it can be concluded that it is critically important for vendors to strive toward "Discovering Reciprocity" and the NISO Open Discovery Initiative–as means of fairly representing usage and providing equitable access to content for all library users, who do not have the option of selecting which discovery system is available to them. When index records are shared, Libraries can properly evaluate usage to achieve our mission: to provide full access to information for the empowerment of research and learning.

Moving forward as centralized indexing develops

As library technology evolves, future research should investigate discrepancies between SUSHI usage data and that which is provided by vendor websites. In preparation for this paper, the authors found that there is a considerable gap in the literature on this subject since the time of COUNTER and SUSHI plans for implementation (as recent as approximately five years ago). Must the literature wait until more robust systems are built? Even then, the problem may perpetuate as the library market narrows. In fact, as the acquisition of Ex Libris by Proquest led to the joint Central Discovery Index, it advertises support for indices with the exception of EBSCO.

Marshall Breeding (2018) advocates for libraries to make their purchasing decisions independent of package deals; however he notes

the trend is toward bundled products. The real question is whether it is really a consideration of money on the table or the convenience of which metadata will best work in a given library's collection for present and future needs. Indeed, K-State Libraries' case study of collection management with competing EBSCO and Ex Libris metadata transfer, and thereby usage data, has shown the inevitable bias. Today SUSHI, while an information standard, remains to be a relatively emerging practice for broad-scope implementation.

Another standard to consider for its implications is the NISO Open Discovery Initiative (ODI). Horava noted that the same sort of exclusivity that K-State experienced between EBSCO and Ex Libris "goes against the endorsed principles of [ODI] which promotes the importance of resource discoverability across multiple channels to meet patron needs" (2019, 17). Looking at the NISO recommended practice for ODI (2014), both of these vendors are in fact compliant in their core metadata provision, but the metadata for resource formats are not specified by NISO to the extent that Ex Libris and EBSCO use. Thus, the formats vary and the mapping issues occur for the API. NISO is again working on the ODI standard, so perhaps a more extensive set of core metadata terms is in the future.

Library advocacy is not only a voice for politics or movements; it should also happen in the business communication exchanges librarians have with their system and product vendors about enhancing usability and comprehensiveness. There are three ways to advocate in this context: through user interest groups, personal interactions with vendors, and possible consortial pressure. The first of these, user interest groups, can be very active in product development. With concern for K-State's centralized indexing, relevant interest groups are EBSCO Information Services Strategic Summit and Ex Libris Users of North America (ELUNA). These interest groups are optimal because the face-to-face meeting and ongoing organization is arranged by the vendors to openly request feedback and share solutions by the vendor's collaboration with the user community. K-State librarians regularly attend the ELUNA conference and their Ex Libris regional group as well as other conferences where they interact with the vendors. The more conferences that are attended and more personal interactions with vendors can lead to a greater level of understanding and awareness of institutional needs, as well as opportunities to convey

local feedback from patrons. Librarians may also choose to organize within their consortia to pressure the vendors by making a larger impact statement and call to action. Together, libraries wield greater purchasing power: consortia have made this their course of business. However, the calls for vendors to cooperate on centralized indexing and metadata are made, libraries literally can no longer afford to be quiet about it.

Disclosure — No potential conflict of interest is reported by the authors.

References

Branch, Denise, Dave Hovenden, and Jodi Upton. 2019. "A Whirlwind of Change Data. How Libraries and Vendors Translate Data into Discovery and Access." *Serials Review* 45 (3):111–8. doi: 10.1080/00987913.2019.1647330

Breeding, Marshall. 2018. "Index-Based Discovery Services: Current Market Positions and Trends." *Library Technology Reports* 54:8.

CARLI. 2019. "Primo VE and the EBSCO Databases." *Consortium of Academic and Research Libraries in Illinois*, February 19, 2020. https://www.carli.illinois.edu/products-services/i-share/alma/EBSCO-API

Conrad, Lettie Y. 2018. "Exclusive Deals in Scholarly Discovery: How They Hurt Users and Pose Threats to Open Scholarship." *The Scholarly Kitchen*. Accessed February 19, 2020. https://scholarlykitchen.sspnet.org/2018/12/04/exclusive-deals-in-scholarly-discovery-how-they-hurt-users-and-pose-threats-to-open-scholarship/

COUNTER. 2019. "The COUNTER Code of Practice for Release 5 - 8.0 SUSHI for Automated Report Harvesting." Accessed February 23, 2020. https://www.projectcounter.org/code-of-practice-five-sections/8-sushi-automated-report-harvesting/

Deodato, Joseph. 2015. "Evaluating Web-Scale Discovery: A Step-by-Step Guide." *Information Technology and Libraries* 34:2.

EBSCO Information Services. 2019. "EBSCO ODI Conformance Checklists." Accessed February 24, 2020. https://www.ebsco.com/odi-conformance-checklists

Emery, Jill. 2020. "Usage Statistics Discrepancies." *SERIALIST LISTSERV*. Accessed February 14, 2020.

Ex Libris Group. 2020. "Configuring the EBSCO Plug-In." Accessed February 17, 2020. https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/Primo/Back_Office_Guide/110Additional_Primo_Features/Configuring_the_EBSCO_Plug-In

Hanneke, Rosie, and Kelly K. O'Brien. 2016. "Comparison of Three Web-Scale Discovery Services for Health Sciences Research." *Journal of the Medical Library Association* 104 (2):109–117.

Hoeve, Casey D. 2018. "Cultural Memory in Danger: Sustainable Information, Preservation, and Technology in the Humanities: A Theoretical Approach." *Collaborative Librarianship* 10 (2). Accessed February 17, 2020. https:// digitalcommons.du.edu/collaborativelibrarianship/vol10/iss2/6

Hoeve, Casey Daniel. 2019. "Resource Management in a Time of Fiscal Scarcity: Combining Qualitative and Quantitative Assessment for Journal Package Cancellations." *The Serials Librarian* 75 (1–4):42–4. doi: 10.1080/0361526X.2019.1576571.

Horava, Tony. 2019. "Dollars and Decision-Making: The MLA Database and the Problem of Exclusivity." *Technicalities* 39 (2):12–15.

Kennedy, Carol Apollo. 2017. "Optimizing Library Services Tracking E-Journal Perpetual Rights: A Discussion among Publishers, Vendors, and Librarians." *Against the Grain* 29 (4). Accessed July 02, 2020. https://against-the-grain. com/2017/09/v29-4-optimizing- library-services-tracking-e-journal-perpetual-rights-a-discussion-among-publishers-ven- dors-and-librarians/

Krumenaker, Larry. 2001. "A Tempest in a Librarian's Teapot: EBSCO, ProQuest, and Gale Exclusive, and Unique Titles." *Searcher* 9 (7). Accessed July 02, 2020. http://www.infotoday.com/searcher/jul01/krumenaker.htm

Lamothe, Alain R. 2014. "The Importance of Identifying and Accommodating E-Resource Usage Data for the Presence of Outliers." *Information Technology and Libraries* 33 (2): 31. doi: 10.6017/ital.v33i2.5341.

Musser, Linda R., and Barbara M. Coopey. 2016. "Impact of a Discovery System on Interlibrary Loan." *College and Research Libraries* 77 (5):643–53.

National Information Standards Organization. 2014. *NISO RP-19-2014*.

Orbis Cascade. 2020. "EBSCO and Ex Libris." Accessed February 17, 2020. https:// www.orbiscascade.org/ebsco-ex-libris/

Rose, Steve. 2018. "Kansas Universities Could Face Budget Axe after State Supreme Court Ruling." *Kansas City Star*, June 29.

Stovold, Elizabeth. 2017. "Evidence Summary Web-Scale Discovery Services Retrieve Relevant Results in Health Sciences Topics Including MEDLINE Content." *Evidence Based Library and Information Practice* 12 (2):169. doi: 10.18438/B8ZH3R.

SUNY. 2019. "Discovery Working Group: Best Practices and Recommendations." Accessed February 22, 2020. https://slcny.libguides.com/discoverywg/best_practices

Thompson, Jo Linda. 2014. *Implementing Web-Scale Discovery Services: A Practical Guide for Librarians*. Lanham, MD: Rowman & Littlefield Publishers.