Purdue University

Purdue e-Pubs

Charleston Library Conference

Reviewing A&Is and Aggregators in a Large Research Library Collection

Holly Inglis
University of Toronto Libraries, holly.inglis@rotman.utoronto.ca

Weijing Yuan
University of Toronto Libraries, weijing.yuan@utoronto.ca

Cristina Sewerin
University of Toronto Libraries, cris.sewerin@utoronto.ca

Follow this and additional works at: https://docs.lib.purdue.edu/charleston

An indexed, print copy of the Proceedings is also available for purchase at:
http://www.thepress.purdue.edu/series/charleston.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences.

Reviewing A&Is and Aggregators in a Large Research Library Collection

Holly Inglis, University of Toronto Libraries

Weijing Yuan, University of Toronto Libraries

Cristina Sewerin, University of Toronto Libraries

Abstract

To facilitate discovery, libraries have traditionally subscribed to many specialized subject abstracting and indexing databases (A&Is), as well as aggregator packages with A&I function and full-text content. Library collection staff must continue to demonstrate effective and responsible stewardship of the library's acquisitions budget by using evidence to inform collection development decisions. Using COUNTER statistics, title lists, the Gold Rush overlap review tool, and feedback from colleagues, review teams at University of Toronto Libraries analyzed a select list of A&Is and aggregators to confirm if subscription renewal is necessary. Involving staff from various departments resulted in a deeper understanding of database use and assisted in deciding not to renew. The review methodology will be used in future reviews, and analysis tools will be shared for future collection development decisions. Discontinuing subscriptions allows us to redirect funds for new resources.

Introduction

The rapid advancement of information technology in recent decades has changed both the role of traditional library reference resources such as abstract and indexing (A&I) products, and the ways in which they are used. What is the role of A&I databases and aggregator packages in today's world, when discovery technology is relatively mature and widely adopted in academic libraries, and libraries license much full-text content directly from the publishers? To what degree do these products overlap in their coverage? What products continue to serve the needs of faculty and students?

To facilitate discovery, libraries have traditionally subscribed to many specialized subject A&Is and often aggregator packages with A&I function and full-text content. A&I databases were one of the early products to evolve from paper to electronic format. Early A&I databases were mostly discipline specific and provided controlled vocabulary and thesauri that faculty and students relied upon for searching the literature. These databases were key tools for content discovery at the time and were often introduced to students as part of library instruction.

With the rapid advancement of technology and the Internet, A&Is grew quickly. Greater numbers of subject-specific A&I products became available with features such as in-depth indexing. At the same time, comprehensive products such as Web of Science and Scopus were developed, which covered multiple disciplines, provided extensive coverage of years, and indexed additional material types. Furthermore, vendors started creating large products by gradually upgrading some A&Is to full-text aggregator packages, often through "inexpensive" or free upgrades to libraries, or bundling or merging products, or removing concurrent user options. In addition to annual subscription increases that often outpaced the growth of library acquisitions budgets, libraries faced fewer options and higher costs. Aggregator packages grew larger over the years and became very expensive to license. As a result, libraries collected many products with duplicate indexing and duplicate full-text coverage, especially of journals that were directly licensed from publishers.

Meanwhile, user needs and behavior have evolved with new technology and service models. Users today do not rely on the traditional reference tools to the same extent as they did 15 or 20 years ago. For many users, Google and Google Scholar are their starting points when searching for information. In addition, the advent of discovery systems in libraries also expanded the search capabilities immensely.

As the demand for access to new scholarly content continues to grow, University of Toronto Libraries has undertaken a comprehensive review of A&I products, as well as a number of large aggregator packages, in an effort to update our collections and streamline services. The goal was to redirect money spent on

A&Is and aggregating products toward new content in order to build a collection that meets the needs of our users.

Getting Started

At the University of Toronto Libraries, the Collection Development Management Committee (CDMC) is a forum to share collection development–related information among 44 libraries in the system and to develop shared projects that impact the larger system. In early 2015, CDMC formed a working group for our aggregator review of ProQuest Central databases, chaired by a CDMC member, and requested title overlap data from the assessment and metadata librarians. A list of A&I and aggregator databases was made available to all library staff for feedback via a Google Docs spreadsheet. Google Docs was also used for storing the review documents in subject area spreadsheets. Meeting minutes, project overview information, and methodology were kept on the library's staff intranet.

When the A&I reviews began, subjects were divided into groups: Humanities, Social Sciences, Life Sciences, Medical Sciences, and Engineering. One chair was assigned to Life Sciences, Medical Sciences and Engineering and another to Humanities and Social Sciences. The groups started with Humanities and Life Sciences and then other groups were formed for the later reviews.

Getting Staff Involved

As the project was being run via working groups, a call for volunteers was first sent to library managers. Working group chairs also went to the Reference Services Committee and team meetings to talk about the project and encourage participation. Some subject selectors and reference librarians were approached directly. In a few cases, staff were unable to make the time commitment to the group so they were copied on group communications in order to see review results and discussion.

The A&I teams consisted of volunteers in different roles, including librarians, technicians, and managers, working in reference, collection development, scholarly communications, and ITS. This brought a variety of perspectives to the table and allowed us to explore new avenues of information in our decision making, such as looking at troubleshoot tickets for database access or content issues.

Methodology

Uniqueness

For the analysis of the unique content, we reviewed how many titles were actually unique within the database. How many had some unique coverage or were truly unique titles? For titles that we had no other holdings for, were they peer-reviewed? Relevant to a particular discipline or program?

Gold Rush, produced by Colorado Alliance, is an overlap analysis tool for database holdings. It can analyze databases containing full-text or citation only, or both. We began using this tool with our A&I reviews. Prior to using it for the project, title lists from Gold Rush were analyzed against title lists from the vendor to see how up-to-date the lists were and to make sure we were working with current data. We were satisfied with the results and subscribed to the product.

On occasion, it was not possible to use Gold Rush because the vendor was unable to supply a title list including ISSNs. In that case, title lists would be compared using Excel for overlap. If the list was unavailable or out-of-date in Gold Rush but was available from the publisher with ISSNs, Gold Rush updated or added title lists upon request.

As Gold Rush does not compare coverage timelines if there is significant overlap between two databases, we would run a date overlap via Excel. In one case, there was complete date overlap of over 50%. The remaining overlap missed a decade of coverage, primarily in the 1980s. The reviewers checked with liaison librarians and selectors to see if the missing decade would present an issue if it was missing, and it was decided not to renew one of the databases.

Gold Rush only compares journals with ISSNs, therefore the title lists from the vendor had to be reviewed for other publications. It gave additional perspective to see how much of the subscription was books, trade publications, dictionaries, or other materials.

Usage

The reviewers used COUNTER statistics when available. As usage may vary within the subject area and program size, there was no set number required for usage. Where data were available, we looked at degree of use in the last three years, defined by number of clicks and number of record views. For

lower usage, we would check program size or other variables that would affect this and explain the lower numbers.

There have been some challenges with high usage for A&Is with no reported use in instruction, Lib-Guides, or reference by the liaison librarians. In one case, we discovered a social science A&I was being heavily used by a medical science program.

Relevance and Scholarly Value

For the majority of databases, the emphasis on scholarly value was confirming that the materials included were peer-reviewed journals. Some data-bases contained materials such as pamphlets, trade magazines, and e-books that were not of interest at the time of subscription or when they had been added to the subscription.

In some disciplines, such as business or education, trade periodicals were considered relevant and were weighted differently. Trade publications may have limited online access if not available via aggregators or an index, making it more challenging for patrons to find articles, unless they know to search the title directly.

For the majority of titles that were reviewed, we focused on English-language publications. If a database contained a significant list of non-English-language titles, and the titles weren't deemed essential to that field of study, it made chances of renewal less likely.

Noting open access journals was also part of the review. If the databases increasingly add open access content, then we can review the cost of the essential unique titles within the database for replacement. By the time of our review, two of the A&I databases on our review list had become open access.

Functions and Features

This was primarily search function analysis, for example, what is the depth of searching available? If all that is available is a simple keyword search, is it sufficient for user needs? Are there any comparable databases with better searching options?

After the first round of A&I reviews, the chairs real-ized that tracking which A&Is were able to link to our full-text holdings was an important feature to review.

Additional Information

Cost

Cost was included in the review criteria for the A&I reviews. It is useful to calculate cost per usage when the option is available. Cost was not used as a target for cancellation; the uniqueness of titles and usage were more important to that decision.

Staff Feedback

Liaison librarians, reference staff, and selectors were contacted. A standardized e-mail to request information was available to reviewers. Follow-up would occur if the response was limited and further explanation required.

Reviewers checked to see if the database was listed in LibGuides or subject A–Z guides. With selectors, we would also ask for information on why the database was subscribed to in the first place, and what the criteria for renewal were. With many databases, the initial reason for subscription was unclear because it had been a decade (or more) prior to the review. Sometimes the criteria for renewal was "other research universities have it in their collection."

Additional Challenges

As mentioned, challenges included title lists without ISSNs, meaning that they couldn't be used in Gold Rush. We also had vendors who could not provide usage stats. Typically, these were small databases developed by faculty, researchers, or associations.

For A&Is with narrow subject areas, low usage required careful review of program size and research interests. Program changes, or changes in the field, could make a database less relevant than at the time of subscription.

Content in A&Is and aggregators shifted during the review period. We also had challenges with publish-ers repackaging the product while we were reviewing it, and having to decide if we wanted to continue with the replacement instead.

Results

For our first aggregator review, we discontinued subscriptions to 26 of the 30 databases within ProQuest Central. This analysis revealed that of the 21,935 titles available from PQC, only 1,809 full-text

journal or periodical titles (8%) were unique and not available to us through other sources.

For the A&I reviews, we have recommended 36 for cancellations so far. It was expected that more of the A&Is would be cancelled, but the reviews showed many to have unique coverage and scope. As usage is declining for a number of titles, a recommendations list has been compiled as priority for review in future years.

The methodology developed for these reviews will be continuously refined and practiced. We will continue using Gold Rush in future reviews as well as in future selection decisions.

Having new librarians and noncollections librarians involved in the project enhanced knowledge throughout the teams. It gave an opportunity for knowledge sharing and informal mentoring. It also built teamwork between staff who ordinarily would not have opportunities to work together, while building capacity and experience with assessment.

Once subscription cancellation decisions were made, there were additional challenges while we worked to make subscription changes. Since some of the

subscriptions started many years ago, we had to face legacy issues of past acquisitions practices. In the early days, libraries invested heavily in A&I products and often licensed some products under a perpetual access model with annual hosting fees, or added an additional year of ownership with each renewal. When we decided that these products were no longer useful, we faced the decision of whether to let them go; that wasn't an easy decision. We also had to work through product bundles that were difficult to unbundle, as some products no longer exist and vendor pricing models also changed.

Conclusions

Libraries' priorities for collection development must evolve with users' needs. While A&I databases and aggregator packages were considered essential resources for many library collections in the past, their roles have changed and their usefulness has decreased, especially in large research libraries with a rich collection of full-text journals and robust discovery services. We learned that many of these products had become redundant. Through this exercise, we were able to remove redundant products, streamline services, and redirect funds toward new content and resource types that our users need.