Spoilt for Choice : A comparison of two aggregations and the use of ScholarlyStats and Ulrichs serials analysis system.

Abstract

The proliferation of online databases supplying aggregations of journal articles to support undergraduate courses means that academic library staff need to exercise care and maintain a watching brief on the performance of these databases. In 2006, the University of Melbourne Information Services decided to cancel one large aggregation database in favour of a competitor, after several years of monitoring coverage content and date ranges.

This paper will define the term aggregation as used in the paper and provide details of both aggregation packages before going on to discuss the rationale behind the evaluation and the procedures and tools used in the evaluation process (Ulrich's Serials Analysis System (USAS) and ScholarlyStats (SS) from MPS Technologies).

In 2006, Jen became the University of Melbourne's CAUL Datasets Co-ordinator. In this capacity, she established a working party to conduct the comparison of the databases. To ensure a representative sample of staff, liaison librarians and collection managers from a wide discipline range were included. There was a perceived need to provide value for money and ensure transparency of decision making and involvement of stakeholders from all relevant disciplines.

The first task was to provide spreadsheets to working party participants. These spreadsheets contained lists of full text titles unique to each package as well as lists of all full-text titles with coverage dates included. USAS was then used to provide subject analysis and to eliminate titles already held online through publisher packages or individual subscriptions.

After consultation with stakeholders, members of the working party returned with lists of titles identified as essential. Jen turned to SS for usage data for these essential titles. It was then possible to recommend purchase of individual title where usage data indicated the necessity, or to point out these titles deemed essential were in fact not used, or little used, during 2006.

The paper will close with the recommendation to continue this type of analysis regularly to ensure our collections reflect the needs of the users while providing financial managers with appropriate data to ensure accountability and value for money.

Introduction.

Electronic resources make up an increasing part of an academic library's collection. There is an explosion of electronic deals including publisher packages, vendor aggregations, evidence based medicine packages and many others. Key resources that should always be monitored are aggregations. An aggregation database, such as Ebsco

Megafile, Expanded Academic or Proquest 5000, is a compilation of journal articles covering an extensive range of disciplines and publishers. Journal content is not cover to cover; for example, editorials are usually omitted. The content is also volatile as publishers can add or withdraw titles at any time. They are, however, a significant resource for academic libraries as they provide access to huge amounts of information and hundreds of journal titles at a reasonable price. They are an especially useful tool for undergraduates.

A significant proportion of the annual budget is spent on these deals; however, budgets have not increased at the same rate. As libraries' budgets come under ever-increasing pressure the need for an organized system to monitor, maintain and support the analysis of usage statistics has become apparent.

It is necessary for librarians to be selective in their choices of new titles and to conduct critical analyses of existing subscriptions. It is important that librarians have the ability to analyse and evaluate electronic resources; consequently, new resources have been developed to analyse and evaluate electronic resources.

Usage Statistics

Academic libraries need to justify the large amounts they spend on collections by demonstrating how they support the key missions of the institution. As digital collections and their usage statistics assumed greater importance within academic libraries, existing usage statistics provided by publishers were seen to be insufficient for evaluation of collections. Journal usage data is valuable for librarians, but gathering the data and presenting it in a useable format has until recently been problematic. Library staff spent many hours poring over statistical output from publishers, trying to establish some common grounds for accurate comparison and evaluation of journal usage. Statistics were often coded in the publisher's own format, and some were only accessible on the publisher's website, and were not available in a downloadable format. Having statistics in a variety of formats makes the task of evaluating and comparing statistics of journal usage extremely complex (if not downright impossible). Publishers, as well as librarians, need consistent measures to demonstrate use. An essential requirement to meet these objectives is an agreed international set of standards and protocols governing the recording and exchange of online usage data. The methods used to measure this use have evolved and continue to do so under the influence of initiatives such as the International Coalition of Library Consortia (ICOLC) guidelines and Project Counter (Counting Online Usage of NeTworked Electronic Resources) codes of practice. (Blecic 2007)

COUNTER

In 2002 COUNTER was launched. This initiative is an example of an international co-operative project involving vendors, librarians and intermediaries such as subscription agents. The international advisory board of COUNTER includes experts from the library, publishing and intermediary environments. Examples include Richard Gedye, the chair of

Oxford University Press, Ann Okerson from Yale University Library, and David Sommer from MPS Technologies.

Its aim is to provide a commonly accepted set of definitions, data processing rules and formats so that usage statistics are standardised and comparable across platforms and packages. Reports available for the journals and databases are

- JR1: Journal Report 1: Number of Successful Full-Text Article Requests by Month and Journal
- JR2: Journal Report 2: Turnaways by Month and Journal
- DB1: Database Report 1: Total Searches and Sessions by Month and Database
- DB2: Database Report 2: Turnaways by Month and Database
- DB3: Database Report 3: Total Searches and Sessions by Month and Service JR3: Number of Successful Item Requests and Turnaways by Month, Journal and Page Type
- JR4: Total Searches Run by Month and Service

COUNTER continues to grow and improve. After introducing a code of practice for journals and databases, it introduced a code for electronic books and reference works. It provides a list on its website of COUNTER-compliant vendors. In the latest list of COUNTER-compliant vendors (dated March 2007) 59 publishers and intermediaries are listed as COUNTER compliant in respect of the most common report format. Examples include *Science Direct* from Elsevier, *Science* from the American Association for the Advancement of Science, and *Web of Knowledge* from Thomson Scientific.

Independent auditing of COUNTER usage reports will be required from 2007 onwards. All vendors compliant with a given Code of Practice must have their reports audited within 18 months of their date of compliance. (Counter Bulletin 2006)

COUNTER-compliant usage statistics are provided from the publishers' websites, but need to be collected individually from those websites, each with its own URL, username and password – a time-consuming process; therefore, the administrative costs of individual provider-by-provider downloads is high. It is up to the library staff member to consolidate and analyse reports. "COUNTER-compliant usage statistics are necessary but not sufficient" (Robinson 2006). Now Scholarly Stats from MPS Technologies provides aggregated data with analysis reports.

Enter ScholarlyStats.

ScholarlyStats is a management portal which provides a single point of access for an institution's vendor usage statistics. It collects, standardises and consolidates journal usage data to provide the library administrator with a suite of monthly usage reports through a user-friendly interface. This product eliminates much of the labour intensive activities involved in collecting and consolidating data from a variety of vendors.

The ScholarlyStats service was launched in early 2006, using COUNTER compliant usage statistics from participating publishers. MPS Technologies developed

ScholarlyStats with the assistance of over 50 participating libraries during the 2005 beta phase. The University of Melbourne was one of the beta testers. Subscribing to ScholarlyStats enables libraries to outsource the collection and consolidation of their usage statistics to MPS; freeing up time to make use of the reports in making collection management decisions. Martha Sedgwick, ScholarlyStats Project Manager, says

> Librarians around the world have told us about the problems and frustrations involved with the collection and consolidation of their usage statistics. Currently, many librarians collect, reformat and merge individual statistics reports every month from a large number of publishers and vendors. Collection and consolidation of these reports wastes large amounts of time and can be very frustrating. (Sedgwick 2006)

ScholarlyStats draws comparisons and allows for analysis between statistics in 1 of 4 COUNTER compliant formats:
- Journal Report 1 – full-text article requests
- Database Report 1 – searches and sessions by database
- Database Report 2 – turnaways
- Database Report 3 – searches and sessions by service

When ScholarlyStats was launched it included 6 Dashboard Reports providing a clear view of trends in usage data:
- Dashboard Report 1 – total number of journals
- Dashboard Report 2 – total use by platform
- Dashboard Report 3 – average use by platform
- Dashboard Report 4 - top use journals
- Dashboard Report 5 – low use journals
- Dashboard Report 6 – zero use journals

ScholarlyStats has an extensive online help facility, which includes suggestions from the user community on use of the reports. MPS staff are also very responsive to feedback from customers, continuing to refine and improve on an already excellent service. For example, in May 2006 MPS developed some new functionality in response to customer feedback. New Dashboard Reports include top 50 journals across platforms, summarising a library's top 50 most used journals across all platforms based on the total use for the year to date, and Low Use Journals (excluding zeros) which details those journals with 10 full-text article requests and fewer for the current year. This report provides the opportunity to review low use titles. In December 2006, MPS Technologies won the Best Library Product of 2006 at the International Information Industries award.

ScholarlyStats is now able to transfer journal usage data into third party systems, including Electronic Resource Management (ERM) modules, using the XML SUSHI protocol. SUSHI (Scholarly Usage Statistics Harvesting Initiative) is a NISO (National Information Standards Organization) led project which enables suppliers to provide usage statistics in a standard COUNTER XML format. The SUSHI initiative is seeking a standard model for machine to machine automation of statistics harvesting. SUSHI had

its first successful test in November 2005, and has partnered with many journal suppliers and vendors, including MPS Technologies.

Delivery of ScholarlyStats reports using this mechanism means that library customers can merge the journal use data of 42 major e-resource platforms, covering over 450 databases, over 70,000 journals, and over 13,500 publishers, with other information contained within these ERM systems, for example subscription information. SUSHI allows completely automated request and delivery of usage reports, saving time and resources for the library. It also allows librarians to see how these resources are used at a granular level, including, for example, cost-per-use of specific titles within a package. (ScholarlyStats 2006)

USAS

Ulrich's Serials Analysis System is a tool for librarians to identify, analyse, evaluate and create reports about their institutions' print and electronic serials holdings. Equally importantly, this system also allows the creation of lists of serial titles the library does *not* hold. It is possible, therefore, to compare an aggregated database the library does hold with one they may be considering for purchase. The detailed bibliographic and publisher data behind Ulrich's Serials Analysis System is updated on an ongoing basis and is refreshed weekly.

In brief, use of Ulrich's Serials Analysis System allows librarians to

- Identify gaps and redundancies in a serials collection.
- Compare your collection against the Ulrich's universe of serials or a select core of titles, groups of peer libraries, or a consortium's catalog.
- Evaluate print and electronic holdings using industry benchmarks and quantitative data.
- Analyze the offerings of top serials publishers and other vendors. (http://www.ulrichsweb.com/ulrichsweb/analysis/default.asp?navPage=4&)

A number of different collection-specific and publisher-specific reports can be generated using Ulrich's Serials Analysis System, for example, peer-reviewed titles, or electronic-only titles. For example, it is possible to create and share reports based on subject area, country of publication, language of publication, or publisher.
Individual reports about current serials subscriptions, Comparison & Overlap Reports, Aggregator Comparison Reports, Top Publisher Reports and other options allow customized analysis to meet an institution's individual needs. The Top Publishers reports module within Ulrich's Serials Analysis System provides comparative information on the offerings of major serials publishers in more than 900 Ulrich's Subject classifications.

The Comparison & Overlap Reports feature allows comparison of any two files, or any number of files to determine gaps and redundancies within a collection. It is also possible to compare aggregator or publisher holdings using these. Of particular importance is the facility to compare the full text coverage of different aggregations.

All Ulrich's Serials Analysis System reports can be downloaded in spreadsheet-ready formats. Librray staff can create downloadable reports about a serials collection by subject category and key criteria. Using a plain-text file uploaded in one easy step, Ulrich's Serials Analysis System creates reports based on a collection that librarians can tailor, save, and share.

As previously mentioned, it is possible to compare an aggregated database the library does hold with one they may be considering for purchase, which is what we did in the following project.

The Project

At the University of Melbourne it was decided to investigate canceling an aggregated database (database A) with a view to purchasing a new one on the market (database B). This decision was taken because product A was considerably more expensive than its competitors and there was an impression that this price differential was unjustified in terms of content or interface.

Jen McMeekin established and directed a working party. She provided background information to the members of the working party, explaining the rationale behind the project, the timeline (and how it was arrived at), the procedures established to conduct the evaluation, and the expected benefits and outputs. Cost comparisons of the databases were also provided.

The project was initiated by defining the criteria involved in evaluating the two products under investigation. It was essential to undertake the comparison at a time that fitted into the timetables of the information librarians: after the first rush of information literacy classes at the start of the year and before the rush to assist students to complete their assignments and hand them in. That ensured that the information librarians were able to focus on the project, without their attention being divided. It was, of course, also necessary to meet the deadline of database renewals with vendors, so it was necessary to report back to the responsible committee (Library Resources and Services Advisory Group) to sign off on the recommendation of the working party in time to meet that deadline.

Working Party Membership

Information Services Division at the University of Melbourne comprises a number of programs. Only the two programs with Library responsibilities were involved in this exercise. These were Information Management (IM), of which the Subscriptions Team is a component, and Information and Education Services (IES), which includes the information librarians and those with collection management responsibilities. Team leaders of each discipline were asked to participate or send a representative. The working party that resulted consisted of a reasonable representation of reference librarians responsible across subject areas. These staff members were committed to their users and therefore were motivated and involved. They were very pleased to be involved, especially

as background work was prepared for them and the exercise did not then impose too much on their very busy and demanding duties.

Prior to the initial meeting, after sending out a call for expressions of interest, Jen used Aggregator Comparison Reports from USAS to compile a number of lists. One of the key outcomes was to identify titles unique to each aggregation and not available elsewhere electronically. Two vital lists provided the unique titles held within each aggregation, subdivided into subject areas. These lists were then run against title lists for publisher journal packages such as Springerlink, Blackwell Synergy and Wiley Interscience, held by the University of Melbourne. This exercise narrowed the number of unique titles down further. Finally the lists were checked against individual journal titles subscribed to electronically. These lists of unique titles were then supplied to members of the working party. Importantly, USAS classified these titles into subject areas allowing these to be distributed to the reference librarians involved in these areas.

The next step was for these librarians to take the subject based lists back to the users and staff. From these lists they identified titles considered essential by both the users and other library staff. The faculties needed to know what they were going to lose and which titles they would gain. (One package has significantly more titles while the other has deeper backfiles. Making the tradeoffs is not easy).

A preference was indicated for the new aggregator (database B) A number of factors contributed to this decision, including relative cost, but a significant factor was the usage statistics provided by ScholarlyStats. Not every journal title in the discontinued database was available in the new subscription, and several academics expressed concern about key journal titles that would be lost with the cancellation of database A. Jen then used ScholarlyStats to look at usage of these titles during 2006 High use titles were identified as being desirable to purchase. Fortunately the number was manageable.

Using the ScholarlyStats *Journal Report 1 – full-text article requests* report for 2006 it was possible to give detailed usage information. This proved very interesting in that a number of perceived "essential" titles had been used very little or, in some cases, not at all. The concerns of the academics were allayed to a certain extent when they were shown the usage statistics, showing low and, in several cases, no use over the year for those titles. We can now follow up with an analysis of enhanced or additional titles on top of the base package to see if these contain the titles lost from in the move from aggregator A.

Aggregator B has proved to be the most highly used of all our databases.

Conclusion

More than ever, university libraries require tools that provide sophisticated data to analyse journal usage. Fortunately, there is now much more "hard data" available to inform collection management decision making. Analyzing and understanding usage statistics can support decision making in a number of ways, for example, it is possible to

assess the value of individual titles or packages with usage data as one metric. Usage data can demonstrate value for money, measure the cost effectiveness of online content purchases, and make budgeting decisions based upon demonstrated use. It can explore usage at the institution and show which disciplines are using online journals and highlight the value that online access provides to users. (Conyers 2006) In particular, Ulrich's Serials Analysis System is a powerful tool collection management within individual libraries and over local, regional and national collections. (Keller & Martinsen, 2006)

Jen was new to the process of carrying out such analyses and so it became a learning experience as well as a real life exercise. As such it opened up a new range of activities and possible projects. The comparison of aggregations was a worthwhile activity. The process of reviewing our collections is continuous and we will be reviewing the outcome of this project prior to the receipt of the renewal notice. This process of review helps to maintain transparency and accountability to our stakeholders. We will continue to use evaluative tools for future analyses to ensure an effective use of the materials vote (budget) and to provide the users with what they really want. The University's federated search engine will form part of our suite of tools. We will load our holdings into USAS so that we can do complete comparisons through USAS including easily identifying individual subscriptions to electronic journals that may be duplicated. As well as this, it will be possible to provide information about the academic quality of our electronic resources using this information provided by the product. Combining information provided by USAS with the tailored usage statistics from ScholarlyStats enables collection management decisions to be made, based on usage and quality aspects such as impact factors. Library staff at the University of Melbourne are able to provide meaningful reports to collection colleagues within the library.

Bibliography

Blecic, D. 2007, Fiscella, J.B and Wiberley, Stephen, 'Measurement of Use of Electronic Resources: Advances in Use Statistics and Innovations in Resource Functionality', *College & Research Libraries. v* .68, no. 1, pp.26-44

COUNTER Counting Online Usage of Networked Electronic Resources 2006, *Member Bulletin*, August 2006

COUNTER 2006, *COUNTER Filter*, viewed September 2006, <http://www.projectcounter.org/news.html#research>.

Keller, A. and Martinsen, D. 2006, 'Serials rationalization at Oxford University Library Services using Ulrich's Serials Analysis System', *Serials, v.*.18, no. 1, pp. 38-44.

Robinson, M 2006, 'ScholarlyStats - the perfect accompaniment to SUSHI', *EPS Insights*, 23rd February, viewed 28 September 2006, <https://www.scholarlystats.com/sstats/EPS%20-%20ScholarlyStats.pdf>.

Sedgwick, M 2006, *MPS Technologies Responds to Library Needs with Launch of ScholarlyStats*, viewed 28 September 2006, <http://www.mpstechnologies.com/pdfs/06.05.05.pdf>.