

New Perspectives in Collection Assessment: flexible, customized delivery of collection evaluation data

Cory Tucker

Continuing Resources & Collections Department, University of Nevada, Las Vegas Libraries, Las Vegas, United States.

E-mail address: cory.tucker@unlv.edu

Reeta Sinha

Academic Sales, Springer Nature, New York, United States.

E-mail address: reeta.sinha@springernature.com



Copyright © 2018 by Cory Tucker, Reeta Sinha. This work is made available under the terms of the Creative Commons Attribution 4.0 International License:

<http://creativecommons.org/licenses/by/4.0>

Abstract:

As collection budgets for libraries are decreasing or remaining flat, libraries have invested much time and effort into the assessment of usage and cost data of library materials. Traditionally, usage data have been gathered and distributed by collection development and management departments which create cost and use reports for individual subject librarians and liaisons. However, this top-down approach to providing collection data is time-consuming and many libraries do not have enough staff or time to provide these reports.

This paper presents a new perspective on collection assessment where collection evaluation and assessment is placed into the hands of subject librarians and liaisons. New technologies and collaborations with vendors and publishers mean collection development departments no longer need to collate and create individual cost and use reports for subject liaisons to review. Software such as Tableau® and customized collection evaluation reports provided by vendors and publishers allow focused, individualized data to be delivered to subject librarians. This shift to streamlined delivery of data is more efficient and cost-effective, making the assessment process easier for librarians/liaisons and the Library. The paper will describe delivery of assessment data via Tableau dashboards and provide examples of vendor/publisher data including, turnaway reports, interlibrary loan reports, comparison reports within and between institutions, comprehensive publisher/vendor account reviews and COUNTER reports.

Keywords: collections, usage, data, evaluation and assessment

Introduction

Over the past several years, libraries have faced decreasing or flat collection budgets. Due to the fact that collection budgets comprise a large percentage of a library's budget, libraries have invested staff time and effort into gathering cost and usage data and it is important to carefully analyze all resources (Fry, 2013). Evaluation of a collection is important because it allows a library to analyze how the resources are being used and can assist the library in shifting money to resources that may have better usage (Fleming-May, 2010).

Traditionally, usage data have been gathered and distributed by collection development departments which create cost and use reports for individual subject librarians and liaisons. However, this top-down approach to providing collection data is time-consuming and many libraries do not have enough staff or time to provide these reports. By adopting new technologies and collaborating with vendors, libraries can more efficiently deliver focused, individualized data to subject librarians.

UNLV Libraries

The UNLV Libraries serve as an example of how collection evaluation data have been collected and disseminated in US academic libraries over the past two decades, since the shift from print-based collections.

Currently, the UNLV Libraries have over 1.2 million monographs, 300 databases and 27,000 journal titles. The collection is overseen by the Continuing Resources & Collections (CRaC) department, which is in the division of Collections, Acquisitions and Discovery. CRaC collaborates with fourteen Liaison Librarians, who are responsible for collection development, reference and instruction for various subject areas. In order to assist the Liaison Librarians with their collection development responsibilities, CRaC provides access to a variety of usage data and usage reports. Usage data are harvested, either by email or logging in to a vendor's site, each quarter, after the 15th of the month. In most cases, usage is delivered via industry-standardized reports. This process is very time-consuming and involves one full-time staff member and a library student worker. The data are then entered into an *Excel* spread sheet and the spread sheets are organized by vendor/publisher. Search statistics are compiled for each fiscal year and divided by the annual cost to give a cost-per-search value or cost-per-full-text-view figures. Harvested data include:

- Vendor
- Title of Electronic Resource
- Monthly searches and full-text views
- Total searches and full-text views
- Annual cost of product
- Cost per search and full-text views (calculated from above data)

Once the data are harvested, staff create a variety of reports and benchmark usage by type of resource, including abstract and index, full-text database, ebooks, and journals. The benchmarking also includes inflation rate and increase/decrease of usage each year. The usage data are placed on a shared drive that is accessible to each Liaison Librarian. Usage reports are also provided to both Liaison Librarians and the Collections Committee.

The UNLV Libraries also collect usage statistics for individual ebooks and streaming media, interlibrary loan (ILL) reports and turnaway statistics, which are provided by vendors. These statistics are also made available to librarians and the Collections Committee on the Libraries' shared drive. Similar to other libraries, UNLV Libraries also gather institution specific information for UNLV faculty including sources and titles cited (Wical, 2015), number of publications in a journal, advisory board participation and journal metrics (Gao, 2016).

Vendor-Supplied Usage Data

As mentioned in the preceding section, the UNLV Libraries, like many libraries, rely on vendor-supplied usage data to evaluate electronic resources such as electronic journals, ebooks and databases. Until the mid-1990s, identifying which titles in a library's collection were used and how often was made possible by analysing statistics available locally. Whether manually collected - shelving counts, for example, or collected by library systems, such as circulation data, libraries were able to evaluate collection use internally and independently. Acquisition records combined with circulation or other usage statistics provided a 'cost-per-use' ratio. Institution-specific subjective parameters could also be applied to calculate the perceived value of one journal over another. A journal with a faculty member on its editorial board, for example, or with a subject scope that was a core research area of the institution – that journal may be of more value to the collection than others, even if use was not as high. ILL statistics provided some measure of demand – journals or books frequently requested coupled with the cost to fill the request could be analyzed and lead to decisions to add (or not) those titles to the collection.

As library collections transitioned from primarily print collections to mostly online, web-based resources hosted by content-providers, libraries began to rely more and more on vendors, publishers and/or other third parties for their collections' usage statistics (Baker & Read, 2008). Metrics such as the number of searches and the number of article or chapter downloads supplemented or, in some instances, replaced circulation statistics. In the early years, however, as print resources went online, different publishers or content-providers handled use differently – or may not have the provided use data at all. When data were available, comparing use from one resource to another was difficult at best – the definition of metrics such as 'views' or 'downloads' might vary from one publisher to the next. Services such as COUNTER (<http://www.projectcounter.org/about/>) emerged to establish industry-wide standards for reporting usage, making it possible for libraries to retrieve and compare usage data across different content-providers. As described earlier, the UNLV Libraries retrieves COUNTER reports on a quarterly basis to assist librarians with their collection evaluation efforts.

Libraries may also avail of non-COUNTER reports and information that go behind the numbers, so to speak. Springer Nature, for example, provides its customers annual Account Reviews designed to analyze longitudinal usage of subscribed eJournals, eBooks and/or databases. Account Reviews also include web analytics for Springer Nature platforms, institutional authorship reports and custom reports as requested. In addition to spread sheets detailing use by title, subject, month, year, copyright year and other parameters, data are summarized in charts and tables for consumption by those who just want an overview, leaving in-depth analysis to others.

More recently, as the realm of scholarly publishing has moved beyond published literature, alternative metrics or altmetrics have emerged as a measure of use and/or impact of online articles and eBooks. Mentions in social media, the news, or blogs are recorded more

immediately than monthly usage reports, for example. A platform such as *Bookmetrix* (©Springer Nature) displays the number of downloads, readers, citations and mentions in social media for each eBook published.

However, even as these usage reports were standardized and new metrics developed, the issue of collating use data and presenting the data to librarians by subject discipline, across all platforms, formats and content providers remained. The protocol SUSHI – Standardized Usage Statistics Harvesting Initiative was introduced to replace the time-consuming task of retrieving COUNTER reports manually for each publisher. Not all publishers supported SUSHI, however. Recognizing these hurdles, third-party vendors offered tools and/or services to collate eJournal use data across publishers and products – an early example was *ScholarlyStats* followed by cost-per-use analyses provided by subscription agents. These services excluded databases for the most part, however, and any publisher-direct subscriptions a library might have. The end-product was generally a file of *Excel* reports which might be re-formatted or otherwise manipulated to share with librarians. Discovery services and/or link-resolver vendors have developed collection analysis tools to facilitate the collection of use for all products from all publishers. Examples include ProQuest's *Intota™* and OCLC's *WorldShare Collection Evaluation*. To obtain the most reliable usage data each month, however, or annually, most often libraries continue to rely on COUNTER reports.

The Next Level - UNLV Libraries

Due to the fact that the gathering of usage statistics is extremely time-consuming and difficult to customize and deliver data to each Liaison Librarian, UNLV Libraries have strategically discussed the use of new technologies to improve efficiency of gathering usage statistics and help with customization of the data. The first strategy was to focus on the library marketplace to see if there are new products that could provide cost-effective options for gathering usage statistics. After a review of the marketplace, UNLV Libraries are using *RedLink®* (<https://redlink.com/>) which is a new subscription service that provides libraries a simple solution for gathering both COUNTER reports and non-COUNTER use statistics. *RedLink* involves a simple set-up and once an account is established, the service automatically harvests COUNTER statistics, including turnaway reports. In addition, *RedLink* has standardized reports to provide cost-per-use, cost-per-full-text, usage by subject area, etc. In short, *RedLink* allows libraries to save an enormous amount of staff time for harvesting data and creating reports.

In addition, UNLV Libraries are making use of the *Tableau*. *Tableau* is software that is an online analytics platform that allows users to create dashboards that make data or statistics more customizable and allows a user to create a variety of visualizations for the data or statistics. *Tableau* enables organizations to place a large volume of data in one place and provides the opportunity for users to use the dashboard to locate more specific data, depending on their needs. As stated previously, the UNLV Libraries gather a large amount of data and it is currently accessible on a shared drive. The data are not easily customized and it would take an enormous amount of time to put together reports for each Liaison Librarian. *Tableau* allows the Libraries to pull together usage statistics and create a dashboard to which each Liaison Librarian can log on and create usage reports for their specific subject area. For example, the Libraries created a dashboard for individual electronic book purchases and usage so. Now, each Liaison Librarian can create a report for their Library of Congress subject areas, using the Dashboard.

Trends for the future

Whether libraries analyze COUNTER and/or non-COUNTER reports and data, or incorporate newer alternative metrics into their collection evaluation efforts, these are all, still, measures of 'use', with some additional information, such as summary presentations or immediate user reactions. Web analytics provide information about how users reach online resources or the devices on which they access the library's online resources. However, just as the UNLV Libraries are moving to the next level in the collation and dissemination of collection evaluation data to librarians, publishers and vendors are already fielding requests from librarians for data and reports that go beyond cost-per-use. It's no longer enough to count downloads or the number of searches, or visits to content platforms. These data answered the 'what' and 'how many' questions. In an effort to understand user behaviour while still required to demonstrate return on investment to administrators, libraries are now also asking why some resources are used and others are not, how collections are used and what happened next? There is also an effort to understand user behaviour that is product-specific – eJournals compared to eBooks, for example, and variations related to format.

Data and information publishers and vendors are being asked or may be asked in the future to provide include but are not limited to:

- Return on Investment (ROI) analyses – a retrospective view of resources used, for example, the percentage of eBook titles used over several years (Lener & Moulton 2017).
- Data related to devices used to access online content – this has implications for both publishers and libraries in terms of user interface design and technical support.
- Top search terms and/or journals with the highest turnaways – what do these data say about research trends at an institution or its curriculum needs? Are library liaisons aware of these data and do collection development strategies align appropriately?
- Platform (web) analytics – referrer source analyses can help evaluate how users reach publisher platforms and user behaviour once there or whether some referring sources are more effective than others. It is worth noting that while these data are useful, drawing definitive conclusions may not always be possible. For example, if users spend less time than some arbitrary average, does it mean they found what they needed quickly or gave up too soon?
- More granular web analytics – how online resources are used. For example, timestamps for journal and eBook use as libraries try to understand when resources are used, how often and for how long. Libraries are now interested in patterns of reader behavior rather than aggregated usage or IPs showing high use.
- Utilization of different functionality or tools – is the user satisfied with a few chapters of an eBook or is the full eBook downloaded? If the full book is downloaded, when, during the same session or on a return visit?
- Demand for other services provided by the publisher – what other actions does the use of electronic resources trigger? For example, will reading an online book trigger a request for a print copy? To answer this, the number of orders for *MyCopy* titles - Springer Nature's eBook print on demand service for institutions that have purchased eBook collections is tracked.

Conclusion

As libraries face decreasing or flat budgets, analysis of usage statistics is becoming more important than ever to help make valid and reliable decisions on a variety of collection management decisions, including serials retention and budget allocation. Usage statistics are extremely valuable in the decision making process, however, the gathering and analysis of usage statistics is very time-consuming. To help keep the gathering of usage statistics more effective and efficient, libraries are encouraged to make use of both current technology and partnerships with vendors. Publishers, content-providers, library-publishers, library software service vendors and libraries must continue to work together to bring collection evaluation to the librarian's desktop, to those professionals who have subject expertise and who work with the end-users. On their part, libraries need to reach out and ask their vendors what's possible – there may be data or metrics not available to libraries that can assist in collection evaluation.

Publishers and other library vendors, whether developing content or providing collection evaluation tools, need to be open to and proactively solicit feedback from libraries, so that the reports and tools they develop are in response to their customers' needs. These data ultimately impact the library's decision whether to subscribe to or purchase their products, or not.

References

- Baker, G. & Read, E. J. (2008). Vendor-supplied Usage Data for Electronic Resources: A Survey of Academic Libraries. *Learned Publishing*, 21.1, 48-57.
- Fleming-May, R. (2010). Chapter 4: Practitioner Responses on the Collection and Use of Usage Statistics. *Library Technology Reports*, 46.6, 28-34.
- Fry, A. (2013). A Hybrid Model for Managing Standard Usage Data: Principles for e-Resource Statistics Workflows. *Serials Review*, 39.1, 21-28.
- Gao, W. (2016). Beyond Journal Impact and Usage Statistics: Using Citation Analysis for Collection Development. *Serials Librarian*, 70.1-4, 121-141.
- Lener, E. F. & Moulton, M. (2017). Taking the Long View: A Case Study of e-Book Usage at a Comprehensive Research University. Charleston Conference; Charleston, SC. Poster presentation.
<https://2017charlestonconference.sched.com/mobile/#session:d15d1c810572a1ba1b36ef6491e4db104>
- Wical, S. (2015). Combining Citation Studies and Usage Statistics to Build a Stronger Collection. *Library Resources & Technical Services*, 59.1, 33-42.