Rolling the Dice and Playing With Numbers: Statistical Realities and Responses

Natasha Cooper
Syracuse University, nacoop01@syr.edu

Kimberly Nolan
State University of New York Upstate Medical University, nolanki@upstate.edu

Mike Poulin
Colgate University, mpoulin@colgate.edu

Nancy Turner
Temple University, tuf56901@temple.ed

Rolling the Dice and Playing With Numbers: Statistical Realities and Responses

Natasha Cooper, Bibliographer, Social Sciences and Interdisciplinary Resources, Syracuse University

Kimberly Nolan, Information Resources Manager, State University of New York Upstate Medical University

Mike Poulin, Head of Collection Management, Professor, Colgate University

Nancy Turner, Assessment and Organizational Performance Librarian, Temple University

Abstract

Librarians invest time gathering numbers about collections for various entities, such as accrediting groups and organizations, as well as their user communities. Gathering collections-related statistics regarding the numbers of things our libraries subscribe to or purchase, as well as the items our users use, often requires a significant investment in time. Definitions can be difficult to apply, and some questions do not seem to reflect our current reality or demonstrate value. The authors explore the challenges of annually gathering and recording collections-related statistics and offer suggestions for improving the process.

Statistical Realities and Responses

Librarians are asked to gather information about collections for reports to various entities including administrators, accrediting bodies, and users. Librarians gathering this information face a number of challenges which include, but are not limited to, determining what data to gather, understanding variations in numbers, understanding the metric definition, using and keeping track of documentation, staff turnover, having enough time, and more. Additional information about these challenges, as well as possible responses, is presented in the following sections.

Issues With Providing Statistics to Reporting Agencies

Mike Poulin, Head of Collection Management, Colgate University Libraries, points to challenges in e-book counting, such as the instructions to submit e-book and usage counts to the Integrated Postsecondary Education Data System (IPEDS). Local cataloging practice makes it difficult to identify e-books, particularly when dual format records are used.

Counting e-book titles available via a discovery service is challenging because for many of the resources being searched, there are no record counts in the administrative module—one just enables the service for search. Others have record counts and can be deduped against other holdings. Pay-per-view titles are to be removed, but it is almost impossible to identify which were purchased versus which are available for purchase. It is also unclear why one would want to remove these records, since it is content that is available for use by our patrons.

For e-book use counts, librarians are instructed to use the BR1 report (book accessed), which is seldom available, but if unavailable to use the BR2 report (chapters accessed). This is essentially the equivalent of combining apples and oranges, since BR1s are access to a title, and BR2s are access to a chapter or section. Since different publishers, for example ProQuest and JSTOR, report the BR2 data differently: For ProQuest, one use equals a page view, and for JSTOR, one use equals a chapter view, so the data is skewed depending on the user preference of platform. As an example, last year at Colgate University Libraries, ebrary use dropped, and JSTOR e-book use increased for a total drop in usage. If one applies the Society of College, National, and University Libraries (SCONUL) multiplier of 5.4 pages per chapter, our use significantly increased. One hopes that the next update of the COUNTER usage statistics will improve comparability between platforms.

A librarian could spend days or even weeks running reports to try to get accurate counts based on the instructions when it is clear that different methodology for counts used between different software platforms results in noncomparable statistics. It is also unclear if the effort involved to do so is producing better counts than a simple methodology.

Reporting to Multiple Entities

Kimberly Nolan, Information Resources Manager, Health Sciences Library, State University of New York Upstate Medical University.

Librarians at Health Sciences Library participate in both academic as well as health sciences reporting. For example, they may prepare reports for the Association of College and Research Libraries, the Association of Academic Health Sciences Libraries (AAHSL), the State University of New York (SUNY) Library Acquisition and Retirement Survey (LARS), and the Integrated Postsecondary Education Data System (IPEDS). Because there are several reporting entities, librarians may need to spend a significant amount of time gathering numbers each year. Staff turnover affects the information gathering process, as new staff assume responsibility for reporting and understanding the ways in which counts were previously done.

It is highly recommended that librarians record the methodology so that someone new can understand what was previously gathered and how. Assigning a survey captain is a good way to make sure your methodology is recorded as well as giving another person insight into each librarian's/department's process. Survey captains can also help keep the process moving and offer support when problems arise.

Additionally, the information gathered through surveys can also be shared with users. The typical library survey provides little appeal to our users. By taking those numbers and adding a few other statistics, we are able to use all that hard work to produce something that is not only more appealing to our users but can go a long way in demonstrating our value. Our Quick Facts (for example: Quick Facts 2015, http://library.upstate.edu/pdf/QuickFacts2015.pdf), which use info graphics and facts of interest, include everything from expenditures to cups of coffee served. This is a great way for librarians to take advantage of

all the hard work put into gathering statistics while also demonstrating what we can and do provide to our users.

Metrics and Assessment

Nancy Turner, Assessment and Organizational Performance Librarian, Temple University Libraries, and member of the Association of College and Research Libraries (ACRL) Library Trends and Statistics Survey Editorial Board.

The collection and reporting of annual statistics can be a burden for even those who love working with metrics. This is particularly true if it means pestering colleagues to generate numbers that

- may or may not demonstrate the value of the library to its parent institution;

- are a meaningful reflection of the 21st-century library's activities and resources;

- or are misleading, confusing, and difficult, if not impossible to gather.

A couple of ongoing discussions in the assessment world related to these issues:

Last year, the Association of Research Libraries (ARL) and the Association of College and Research Libraries (ACRL, ALA) formed a joint advisory task force to suggest changes to the current definitions and instructions accompanying the Integrated Postsecondary Education Data System (IPEDS) Academic Libraries (AL) Component for FY 2015.

As noted previously, the IPEDS instructions for counting e-books originally said to "Count e-books in terms of the number of simultaneous users," a problem if we have a license that has no access restrictions. IPEDS asked libraries to not include open access resources, even if they were in the catalog or the library's discovery system.

The task force made a formal recommendation to IPEDS, and now these definitions are being changed—a good example of how these organizations need to start talking with one another to establish common ground. The examples are also a reflection of how traditional metrics for describing print collections are not always easily translated into the digital world.

Another conversation that took place recently at assessment conferences and through listservs related to whether our traditional metrics were meaningful (Horowitz, et al., 2015). The question was prompted by the publisher of Peterson's college guides requesting the count of a library's microforms. Now what kind of high school student chooses a school based on the library's collection of microfiche?

The informal discussions yielded ideas for potentially, more meaningful metrics, for example:

- number of seats/study rooms/etc. available 24/7 or 24/5;

- availability of self-service check-out;

- library staff to student ratio; and

- outcomes assessment results (correlations to GPA or other).

While these may have relevance to the values and practice of some libraries, not all would be able to generate these numbers or have practical use for them.

Challenges: Consistently Measured and Continuing Relevance

In our data collection practice, libraries face several challenges. We must maintain consistency of measurement over the years in order to be able to see trends and yet stay relevant to changes in libraries: Content types, format types, changing approaches to the acquisition of content, and its discovery in our local online tools

Trend analysis requires that we collect the same metrics year after year, defined in the same way. Title and volume counts are good examples of metrics developed in a print world that don't translate well to the electronic environment, and

while organizations such as the National Information Standards Organization (NISO) have done an excellent job toward codifying these definitions, discerning trends may be like comparing apples and oranges.

Other challenges for counting relate to the increased use of open access resources, discovery layers outside of the catalog (with limited reporting tools), and content made available to users but not technically acquired until it is used.

Suggestions

Looking ahead, we recommend:

- Providing feedback to entities that gather statistics in order to ensure greatest possible relevance for the data being gathered. For example, Project COUNTER participants have invited feedback regarding their work on the COUNTER initiative (see: https://www.projectcounter.org/). The USUS site, which is a community website on library usage (http://www.usus.org.uk/), provides an opportunity to report problems. The LIB-STATS listserv is also a forum for exchange of information about usage practices.

- Encouraging the practice of gathering numbers that are meaningful and that have assessment value, as opposed to solely counting numbers of items.

- Thoughtfully sharing information with user communities.

- Documenting locally and planning for staff change.

- Sharing best practices.

References

ACRL Academic Library Trends and Statistics Survey Editorial Committee (Robert Dugan, chair). (2016). ACRL, ALA and ARL joint task force on IPEDS academic library (AL) component. Retrieved from http://libguides.uwf.edu/c.php?g=291123

ACRL, ALA, and ARL Joint Task Force (2015, July 14). Recommendations to IPEDS—Academic Libraries (AL) Component. Retrieved from https://drive.google.com/file/d/0B3kAuloKaWWTUnk2N25oQmlYT3c/view?usp=sharing

American National Standards Institute. (2013). ANSI/NISO Z39.7-2013, Information services and use: Metrics and statistics for libraries and information providers—Data dictionary. Retrieved from http://z39-7.niso.org/

Association of Academic Health Sciences Libraries (AAHSL). (n.d.). Annual statistics. Retrieved from http://www.aahsl.org/annual-statistics

Association of College and Research Libraries. (n.d.). ACRL 2015-16 academic library trends and statistics survey. Retrieved from https://acrl.countingopinions.com/

Association of Research Libraries (n.d.). ARL statistics & assessment. Retrieved from http://www.arl.org/focus-areas/statistics-assessment; and ARL-ASSESS email list

COUNTER. (n.d.). Retrieved from https://www.projectcounter.org/

Integrated Postsecondary Education System/National Center for Education Statistics. (n.d.). Retrieved from http://nces.ed.gov/ipeds/

JISCM@il (n.d.). LIB-STATS listserv. Retrieved from https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=LIB-STATS

USUS: A Community Website on Usage. (n.d.). Retrieved from http://www.usus.org.uk/

Acknowledgments

Bucknell, T. (2012). Garbage in, gospel out: Twelve reasons why librarians should not accept cost-per-download figures at face value. *Serials Librarian*, *63*, 192–212. http://www.tandfonline.com/doi/abs/10.1080/0361526X.2012.680687

Central New York Library Resources Council (CLRC) Usage Statistics Interest Group. (May 6, 2016). CLRC Headquarters; Syracuse, NY; Presenters and discussants:

- Wendi Ackerman, Deputy Director, Health Sciences Library, SUNY Upstate Medical University;

- Claire Enkosky, Public Services Coordinator, Manlius Library;

- Mike Poulin, Head of Collection Management, Colgate University;

- Tom Keays, Electronic Systems and Resources Librarian, Falcone Library, Le Moyne College;

- Amanda Travis, Member Services Coordinator;

- Scott Warren, Associate Dean for Research and Scholarship, Syracuse University Libraries.

Donovan, G., & Fishel, T. A. (2015). Introducing the new ACRL trends and statistics survey/changes for 2015. *College & Research Libraries News*. *76*(10)*, 540–543.

Horowitz, L. (2016). Meaningful library metrics for marketing. ULS-L listserv post, August 31, 2016.

Horowitz, L., Kinsley, K., Koltay, Z., & Szentkiralyi, Z. (2015). Metrics with meaning: How we can effect change to library assessment metrics used by non-library organizations. Southeastern Library Assessment Conference. Retrieved from http://scholarworks.gsu.edu/cgi/viewcontent.cgi?article=1017&context=southeasternlac