

Purdue University
Purdue e-Pubs

Charleston Library Conference

We're Not So Different, You and I: How Librarians and Publishers Use Statistics in Different Ways to Achieve Similar Goals

Sarah D. Schulman
Springer, Sarah.schulman@springer.com

Alexis Linoski
Texas A&M System Libraries, alinoski@library.tamu.edu

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>.

Sarah D. Schulman and Alexis Linoski, "We're Not So Different, You and I: How Librarians and Publishers Use Statistics in Different Ways to Achieve Similar Goals" (2015). *Proceedings of the Charleston Library Conference*.

<http://dx.doi.org/10.5703/1288284316295>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

We're Not So Different, You and I: How Librarians and Publishers Use Statistics in Different Ways to Achieve Similar Goals

Sarah D. Schulman, Account Development Specialist, Springer

Alexis Linoski, System Libraries Resource Development Director, Texas A&M System Libraries

Abstract

Statistical gathering and analysis plays an important role for librarian and publisher alike as we strive to predict and decipher end user behavior and improve collection development. During this session, hear from a librarian who works with a large university library consortium, and a representative from a major publisher, on how statistics factor into our workflow—and the stories they can tell about how content is being discovered and used. Learn about specific projects and applications from each side, and hear how analyzing end user behaviors helps refine everything from services we provide to institutional decision making. There will also be time for audience members to ask questions and share their own experiences.

Springer

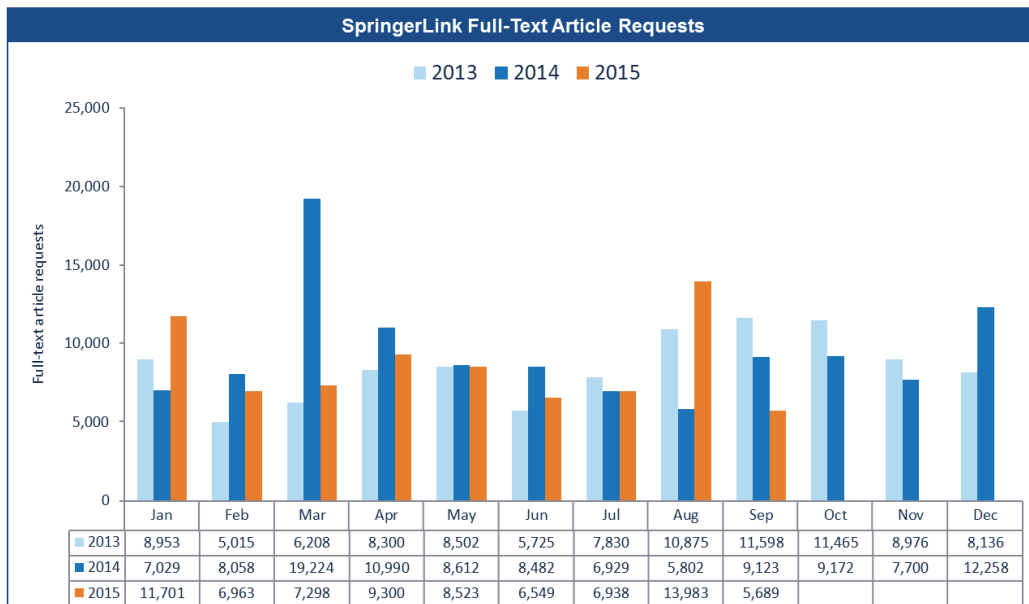
The Account Development (AD) team at Springer collaborates with academic and government library customers to help them attain the best possible value from purchased content. Statistical

gathering and analysis is just one method they use to work toward this goal.

What type of statistics do we look at and why?

1. **COUNTER reports:** On a quarterly basis, each AD team member gathers and records in a central datasheet the

Article downloads by month Jan 2013 – Sept 2015



Source: COUNTER Reporting / BW

Figure 1. This chart compares three years of journal article COUNTER usage on a monthly basis.

English-language packages: chapter denials by eBook package Jan – Sept 2015

eBook Package	Contemporary	SBA
Behavioral Sciences	146	192
Biomedicine & Life Sciences	653	693
Business & Economics	227	88
Chemistry & Materials Science	187	304
Computer Science	10	74
Earth & Environmental Sciences	56	183
Energy	9	
Engineering	406	86
Humanities, Social Sciences & Law	515	316
Mathematics & Statistics	236	633
Medicine	1,205	336
Physics & Astronomy	80	237
Professional & Applied Computing	20	
Total	3,750	3,142

Source: COUNTER Reporting / BW; Springer Book Archive = CY 2004 or older; excluding Protocols

Figure 2. This table shows e-book chapter denials, or turnaways, by subject collection and age (contemporary versus archival).

COUNTER statistics for every “key customer” in our dedicated regions. A “key customer” is defined as one who has purchased or renewed any Springer products that year (including journals, books, and databases). The AD team mostly works with academic institutions with a Carnegie Classification of Master’s Large or higher, although there are some exceptions. In addition to COUNTER statistics, the central datasheet for each region shows how much customers spend per product, the previous year’s usage, and a percentage showing how much COUNTER usage has increased or decreased from the previous year.

Gathering COUNTER statistics on a regular basis and recording them in a central document provides a one-stop method of monitoring usage, spending, and cost per download. Is usage increasing each year the way we want it

to in relation to spending increases? Where is usage decreasing or remaining static? Keeping an eye on this information helps the AD and sales teams ensure that all customers are getting the best ROI and usage.

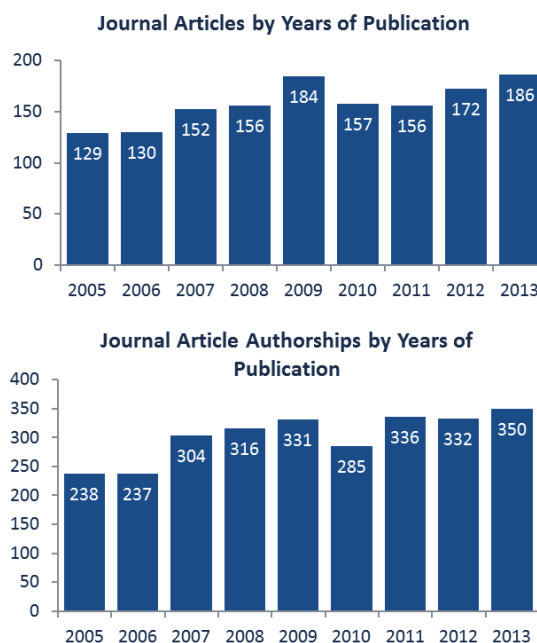
2. **Account reviews:** Throughout the year we visit most (if not all) key customers along with our regional sales counterparts to present account reviews. These detailed PowerPoint presentations show a visual, in-depth analysis of the most current usage and denial statistics per product. Please see Figures 1 and 2 for specific examples.

Performing these account reviews serves a number of purposes. Much like with the central document for quarterly COUNTER statistics, it provides an easy way to compare usage trends per customer on a monthly basis in addition to annually. By looking at the most and least used subject

Journal Articles Published with Springer 2005 to 2013

• Journal Articles Published

- Between 2005 and 2013 researchers from *University X* published 1,422 journal articles in 527 Springer journals.
- Considering multiple authorships, we see altogether 2,729 contributions from researchers from *University X* during that time period.



(Source: Springer Authormapper. Reported numbers are minimum numbers, based on available information about roles, affiliations, and addresses.)

Figure 3. This chart shows the number of journal articles published by authors affiliated with University X from 2005 to 2013.

areas, customers can see which collections might benefit from additional marketing or instruction on campus. Additionally, an account review might reveal problems with the content's accessibility or discoverability which require troubleshooting (either by Springer or the library).

3. **End user behavior analyses:** Usually provided as a complement to the account review, end user behavior analyses look at metrics provided by a licensed service called WebTrek. This product works in real time and can drill down by individual institution to show how end users arrive at the SpringerLink platform with such metrics as:

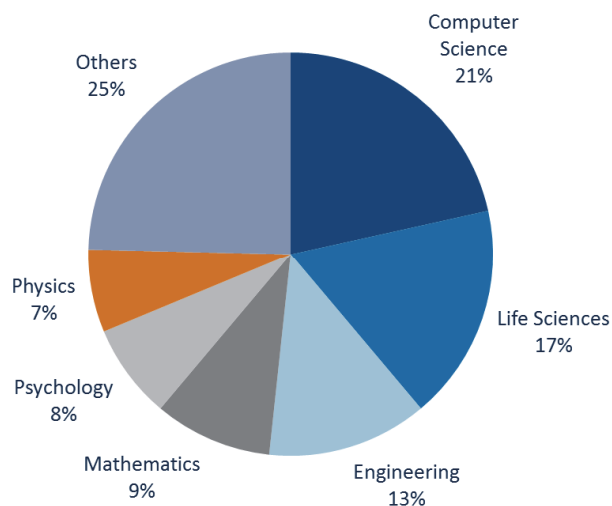
- Entry sources or referrer URLs
- Search phrases. These may be external (for example, what a user typed into Google while on campus

that took them to SpringerLink) or internal (which search terms the end user is entering into SpringerLink's native search box)

- Number of unique and returning visitors, plus the average amount of time they spend on the site and number of pages looked at
- Technology used (for example, how many end users were on a PC or laptop versus mobile device? Which operating systems/web browsers were most used?)
- Most popular visit months, weekdays, or times

There are many more metrics that WebTrek is capable of analyzing, and they can be examined using multiple time frames. Presenting this information provides additional context to the

Book Chapters by Subject Area



(Source: Springer Authormapper. Reported numbers are minimum numbers, based on available information about roles, affiliations, and addresses.)

Figure 4. This chart shows volume of book chapter authorships affiliated with University X by subject area.

COUNTER usage reports and account reviews we show to library customers, and can tell a story about the people behind the numbers. Much like with the account reviews, it may also reveal certain access or discovery issues. For example, if a particular library has a full discovery layer but none of the top 20 referrer URLs are from their link resolver, this may warrant further review.

Authorship analyses: Occasionally and upon request, Springer will analyze the volume of authorship affiliated with a particular institution and present it to the library. These analyses examine journal articles, whole books, and book chapters to provide the total numbers (as a whole and also accounting for multiple contributions) along with a breakdown of authorship by subject area. Please see Figures 3 and 4 for specific examples. These reports can help librarians monitor and support local publishing and research trends. They may also help to facilitate

relationships between faculty and librarians, or even boost usage of homegrown content. From a marketing standpoint, librarians and publishers might collaborate and use local authorship as a “hook” for usage-based end user campaigns.

4. **Miscellaneous:** Customized usage, cost, or other types of analyses may be provided as needed or upon request. For example, a customer may wish to see the cumulative usage of an e-book package purchased in 2010, or a cost per title versus list price analysis. These miscellaneous reports serve a range of functions both internal (for example, help facilitate a sale or renewal) and external (for example, assist in collection assessment projects).

Next Steps: What Do We Do With This Information?

The next steps Springer takes after providing one or all of these analyses vary depending upon the

situation and goals on both sides. For example, during a routine annual “check-in” style of visit, we might just say, “how do you feel about what we just saw? Any concerns or questions?” and leave it there. Where needed, these reports can serve as a springboard to plan usage-boosting initiatives such as trainings, end user marketing campaigns, or reviewing Springer’s discoverability within a school’s online system. Conversations with librarians are invaluable to learn which information is most useful, which in turn helps Springer to inform future goals and best practices.

Texas A&M System Libraries

Why Review Usage Statistics?

Libraries generally review usage statistics to determine if resources *are* being used and, if so, what is the cost per use. Usage statistics can also be used to determine gaps in the collection and to determine *how* the resource is being used.

Usage is not the same for all resources. For databases, journal packages, e-book collections, and streaming media collections, searches or full-text downloads are normally used, with cost per use being determined based on the cost for the database or package. For individual journal subscriptions and e-book purchases, full-text downloads or title usage/circulation numbers are used, with cost per use being determined based on the cost of the subscription or book. The same applies to individual purchases of streaming media.

Almost all libraries rely on COUNTER reports. These standardized reports provide article/full-text downloads, database/collection usage, and title usage, which are used to determine cost per use. They also provide a variety of other reports, including turnaway reports, which show the number of attempts to access material not available through the libraries current subscription or collection. While COUNTER reports provide an “apples to apples” comparison for electronic resources, not all vendors provide COUNTER reports and not all resources are

suitable for COUNTER reports (statistical databases are an example).

What Is Good?

To determine what is “good” usage, libraries generally set a threshold for cost per use, which can vary by resource type. Anything below the threshold is good and anything above the threshold is bad. When setting the threshold, libraries should take into account the average cost of ILL and cost per article from the vendor or other sources (such as the Copyright Clearance Center). The same criteria, either Searches or full-text downloads, should be used for databases and collections. Title usage for these resources may not be as useful because the vendor can add or remove titles at any time. With the advent of Discovery systems, Searches may not be the best criteria to use as Discovery systems can skew those numbers. However, not all resources are included in Discovery systems, so Searches may be a good statistic for those resources.

“Bad” usage statistics can tell us a lot and it’s not always that a resource should be dropped. Before eliminating a resource due to poor usage statistics, consider:

- Are users aware of the resource? Do marketing efforts need tuning?
- Does faculty outreach need to be increased?
- Are there searches but no full-text downloads? If so, perhaps:
 - The resource is not being used appropriately, i.e., a Humanities resource is being used for science research.
 - Users may need instruction.
- Consider online tutorials for distance students.
- If journal subscription usage is low on the vendor platform, is the title included in another resource? If so, what is the usage for that platform?

- The user population is small, i.e., the resource supports a small degree program.

If none of these factors contribute to the low usage numbers, then perhaps the resource should be cancelled.

It should be noted that while purchased e-book titles can't be cancelled, specific title usage or non-usage can provide information for future collection development.

An invaluable collection building tool is the Turnaway Report. Turnaway reports can help inform collection development decisions regarding subscriptions to specific journals or purchase of specific e-book or movie titles (for streaming media collections). Turnaway reports can also help inform decisions regarding resource settings, i.e., should unsubscribed content be hidden from default searches.

There's More

While COUNTER reports are good, many libraries are seeking more information about how their

resources are being used. Non-COUNTER reports are a good place to start, and many vendors offer these reports in addition to COUNTER reports. Two types of reports that stand out are "Top Search Terms" and "Top Referring Sites." Top Search terms can be valuable in learning not only what terms users are searching for, which can give insight into current course assignments, but also *how* they are searching—is it natural language or are they using Boolean operators? Are they broad searches, such as "C+" or very specific searches, such as "students for a democratic society"?

For individual e-book purchases from the same vendor, title circulation and usage can be enlightening. With purchased e-books, simple title usage or circulation is about that is needed to determine if purchased titles are being used. Hold lists and suggestion lists also provide information regarding collection development.

Other reports that might prove interesting are those that give the time of highest usage, number of user accounts created, and length of visit.