# Practitioner Responses on the Collection and Use of Usage Statistics

## Abstract

In addition to the theoretical implications of use and usage, there are practical implications for librarians who interact with use and usage statistics on a daily basis.

# **Usage Statistics Survey**

The results of an informal survey distributed May 4, 2010, to the Electronic Resources in Libraries (ERIL) e-mail list offer a glimpse into how usage statistics are actually being collected and used in libraries. At the time of distribution, there were 1,938 subscribers to ERIL-L and 112 individuals responded to the survey (see appendix). When asked how often (if at all) vendor-provided usage statistics for electronic resources are collected, 45 respondents stated annually; 16, every six months or twice a year; 51, monthly; and 11, quarterly or every three to four months (figure 16). However, these numbers are not as clear-cut as one might assume from first glance. Some respondents noted different time frames for different scenarios; thus, some responses were counted twice. For example, one respondent indicated that database statistics are collected once a year and e-journal statistics are collected twice a year. Moreover, even for those libraries that report collecting statistics on a monthly basis, it still depends on the publisher's or vendor's release schedule. For example, some publishers and vendors release statistics only on a quarterly basis, so monthly collection is not possible.

More than half of the responses included a caveat or further explanation of some sort; caveats ranged from "monthly, but as needed for special reports" to "once a year for annual reports but also at time of renewal for

# How often (if at all) do you collect vendor-provided usage statistics for electronic resources?



Figure 16 Results of question 1 (see appendix ) of the survey.

individual digital resources if under consideration for cancellation" to "usually twice a year, unless a faculty member needs usage stats ASAP." Other responses were not time-frame specific but rather stated "ongoing throughout the year" or the simple yet honest "not very often" or "I try for once a month. It rarely happens."

It is tempting to believe that the 51 respondents who collect statistics monthly outsource at least some of the responsibility through the use of services such as Serials Solutions 360 Counter (figures 17–20) or Scholarly Stats (figures 21–24). However, 32 of the 51 respondents who collect monthly are doing so manually, without the assistance of a third party or without the use of the SUSHI protocol. Of the entire 112 respondents, 41 did indicate that they were investigating, using, or had discontinued a vendor-supplied service. Two respondents quoted budgetary



THE 360 COUNTER ASSESSMENT PROCESS, courtesy of Serials Solutions (http://www.serialssolutions.com/360-counter/)

## Figure 17

Serials Solutions explains its process for 360 COUNTER.

	Titles by Subjected 05/13/2009 12:26:		this Report ]							
🚯 Header 🔻	Metric 💌	🔓 View 💌 🔀 Eb	kport PDF	ort 📄 Save as .	🔘 Hel	p				
Applied Filt	ters: [ Subject: Scie	nces - General 🕅	Month: Jan 2007 to	Mar 2009 😰   Libra	aryCode: DC	<u>s</u> [2]]				
				Year ♥ ♥ŸΣ⊠	20	<u>07</u>	2008		Grand total:	
Subject	Authority Title ເ⇒♥∑⊠	Journal Name ⇔ΥΣ⊠	Provider/Platform ⊜Ƴ∑⊠	Publisher ⇔γΣ⊠	Cost per Use	Uses	Cost per Use	Uses	Cost per Use	Uses
Sciences - General	Ambio	Ambio	JSTOR	Alliance Communications Group	-		0.01	271	0.01	27
		AMBIO: A Journal of the Human Environment	BioOne	Royal Swedish Academy of Sciences	-	-	-	72		7
	American journal of science (1880)	American Journal of Science	HighWire Press	American Journal of Science	-	-	0.04	36	0.04	3
	Annals of the New York Academy of Sciences	Annals of the New York Academy of Sciences	Blackwell Publishing	Blackwell Publishing	0.04	193	0.05	127	0.04	32
	Biographical	Biographical	JSTOR	The Royal						

## Figure 18

Formatting of headers and footers within 360 COUNTER. Courtesy of Serials Solutions.

constraints. One commented that statistics were collected "manually, but only because we do not have the budget to outsource." The second explained: "The quoted cost of Serials Solutions and Scholarly Stats was too high for us," and went on to qualify, "Also, these products can't be used for non-COUNTER-compliant stats."

Not all statistics are treated the same or necessarily collected on the same schedule. One respondent explained: "I try to do monthly reports on some things, calendar year on others, and fiscal year on others. It really depends on the resource as some I need to keep closer track of than others." For libraries that outsource their statistics-gathering duties, time frames may vary as well. One respondent noted that statistics are collected "annually, except for the automated stats from Scholarly Stats we collect monthly. For about 40 vendors, we outsource collection to Scholarly Stats and automatically via SUSHI upload the stats into our III ERM system. For other vendors, we manually collect stats" on an annual basis. Another respondent noted that statistics were collected "all the time, but we're aiming for twice a year. Serials Solutions collects most JR1s for us."

For the already overtaxed librarian, the perceived time needed to implement protocols such as SUSHI plays



## Figure 19

The chart on the left is a graphical representation of titles that received a large percent increase in their usage over the course of a given time period. The one on the right illustrates the fluctuation in uses of titles over a span of months. Courtesy of 360 COUNTER, Serials Solutions.

Journal H	Counter Consolida oldings Repo Metrics Layout	rt - (JR1 and JR	2) (Link to this Report )		
Metric	Placement	Order Clear Default	Formula	Conditional Display	
Uses	Report +	1	+iTotal(Uses)	1(Change(Uses)	Edt
Cost per Use	Report •	2	STotal(Cost per Use)	H %Change(Cost per Use)	Edit
Journal Cost	Metrics pool +	3	HTotal(Journal Cost)	Schange(Journal Cost)	Edit
Turnaways	Metrics pool ·	4	%Total(Turnaways)	Schange(Turnaways)	Edit

#### Figure 20

The advanced metric option in 360 COUNTER gives you the ability to calculate your percent change for the cost and use of a report. Courtesy of Serials Solutions.

a large role as well, even with the use of a third-party service: "We have Scholarly Stats for our nine largest platforms.... I have been wanting to implement SUSHI, but just haven't had the time." Other comments included "I have plans to try to implement SUSHI" and "I'd love to have them [vendor platforms] harvested using SUSHI but who has the time to get IT to do that?" The value of SUSHI may be widely heralded, but its limited use by the vendor and publisher community has certainly not gone unnoticed—"We are attempting to use SUSHI when it works, and are looking forward to its wider adoption. Other than that, it [statistic collection] is manual." Librarians are excited and anxious for the different systems and



#### Figure 21

Scholarly Stats offers a graphic explaining its overall process. Courtesy of Scholarly Stats, a service of Swets.

protocols to work and work well *together*. At present, however, even for the most aggressive of implementers,



#### Figure 22

Scholarly Stats offers another graphic explaining more about its process. Courtesy of Scholarly Stats, a service of Swets.

for clearer results	scholarlysta				
Name Annually Advant Spalar			Marrie S. C. alue	, 11 m · · · · ·	titled including make impact
fairers 1815 Accessity, Materialy of Alberta					
atest Reports					Ratinal tenas
Tur latest tehelafrikata reporta- tomattel (2011), or through a con- tica tena ta doarniaad depending reo here are gathed, locent leng	ere ansledet balane. These reports are analytic to geen or close for pressed op Med 100%, prese und das no mis con motiform side to sport determic contexts. Our reconstration side to an electrical and an electrical to a second that the determined the standards in you care Context I.s. Bittane to the Deport Bitma po- has Edwards by all Albertis, Mar-2008.	eed record. Note the TCP remains as it a	oclarge report will do write ad t	to may take a	State take data taken data taken data taken taken taken metalaka tak metalaka taken metalaka
Concritidated Acquirity	Levengenee	1.00	2387	- 81.5	Drivery Dulloury, Lerone Hermond &
Carrier Andrew Territors ( )	Full-Text Article Responds to Scienced Title, Parliets and Parls	विवन्ध	(1)(1)(1)	Barro	Ephena Askanow have the (Assaults of Well-second side to the fitnesses of interacting standardiate
Development &	Book Title Requests by Title, Harborn and Harth	Re en	Same	10 (194 HT)	app spectaries provide an approximate
Canal Idam & Danial and Amport L	Searches and Second by Holdh and Database	20.00	Suries	Bansie	Querte a
Conversion Decision Report 2	Turnarors he Plerth and Database	TR(L +A)	Sares.	10 (et en:	"I are some away by the
Conservation of Darlads and Resport 3	Searches and Decements Month and Service	30.00	Steen	Bions	sparity of the reports, as well as the strongs and clipter. For the first time toreats
Second Statistics of Reports	Ce actuation	110	63W	- 10.5	company mu station per
Calificant Percet 1	Norther of Tilles by Hattory	Q(2.64.12)	150.72.40	There and	lastile and comprohensies
Latener: Note: 1	Nember of Full Text Article Respects to Platform	Tizmen)	S(1+8)	Tant en:	alextrone haroweak carbonan
Ladding Privet 1	Annyage Line für Flatture	Zawa)	Slove	Eliszant.	
			NUMBER OF STREET	and the second second	
Employed Report #	Teo Unit Assessh by Pattions	RICKAN	S0000	- CLAD KAL	

#### Figure 23

An example of reports available within the Scholarly Stats interface. Courtesy of Scholarly Stats, a service of Swets.



#### Figure 24

An example of a COUNTER-compliant report available within the Scholarly Stats interface. Courtesy of Scholarly Stats, a service of Swets. an uneasy mixture of approaches remains: "We are experimenting with SUSHI now, and hope to employ our ERM stat module (Verde) along with SUSHI for something that would be more robust in the future, but right now everything is pretty piecemeal."

Lest we think that third-party services are the silver bullet, it is important to note that the disadvantages of third-party services emerged as well: "We use Scholarly Stats (Swets) service; however, the downside to that is that stats are approximately two months behind, i.e., vendors do not have to provide COUNTER-compliant stats for 20 days after months end, and therefore Scholarly Stats does not collect until 20 days after that time! So, although their reporting is interesting, particularly for title level stats we can't compile ourselves, we still collect platform manually so that the subject librarians have some sort of up-to-date feedback each month." Another respondent explained: "Right now, we're manually collecting stats. We used to use Scholarly Stats pre-recession. We've found that once we set up our admin accounts, it doesn't take an undue amount of staff time to retrieve them." Clearly, the local library structure and collection size may dictate the cessation of a third-party service or a mixture of third-party and manual collection.

When asked what mechanism is used to collate statistics (e.g., Microsoft Excel, ERM, other), all but nine respondents used Excel either exclusively or in concert with an ERM, XML, SharePoint, Microsoft Access, or other locally developed solution. Of the nine that did not use Excel, Serials Solutions 360 Counter, ERM, and Access and ASP programming were among the tools used to collect and collate the statistics. Clearly, Excel is still at the forefront in terms of statistic collection and collation.

The distribution of usage statistics to internal library personnel occurs in a number of different ways, but e-mail and intranets, protected webpages or shared drives primarily dominate the distribution channels. When asked how distribution occurs, 43 respondents indicated an intranet, protected webpage, or shared drive, and 38 respondents indicated e-mail; 25 respondents explained that they did some combination of sharing via a centralized location (not an ERM or third-party provider) and e-mail upon request. Only three respondents specifically stated that they give librarians access to their ERM or third-party provider.

To determine how collected usage statistics are used, respondents were asked to check all that applied: collection development/management; promotion/public relations, assessment/evaluation; reporting to accrediting or other agency (e.g., ARL); training library staff; and other (figure 25). The clear leaders in this category were collection development/management with 94.6% (106) and assessment/evaluation with 92% (103). Less popular were reporting to accrediting or other agency (e.g., ARL) with 50.9% (57) and promotion/public relations with 42.2%

## How does your institution utilize the electronic resource usage statistics it collects? (Check all that apply)



Figure 25 Results of question 7 of the survey.

(45). Training library staff trailed last with 7.1% (8). The other category, at 14.3% (16) included responses such as these:

- "The data is collected for the purposes above, but frankly it is seldom used."
- "We use stats to demonstrate the library's return on investment by the University. Lately we have had to demonstrate to specific departments the support they get from the library."

Of special note is one response: "I have used them to claim access/credit on rare occasions. If you can show lots of zeros and then several months of consistent use, then it is pretty easy to convince the publisher that there was no access initially." It would be interesting to discover how successful this tactic is.

It is evident from the "other" category that usage statistics are used in budgeting decisions. Ten of the 16 responses cited budget or budget justification as a common use of the usage statistics. One respondent noted, "For our 2010 renewals, we did not renew anything with a cost per download of over \$35." Budgeting might also have been interpreted as being part of collection development/ management or assessment/evaluation. Considering the economic downturn that began in 2008, it is easy to conclude that libraries are looking for ways to justify cancellations. However, the economic downturn has not ultimately changed the fact that libraries are using usage statistics to determine renewals and cancellations. A survey conducted in 2006 found that "the most cited purpose for reviewing and analyzing vendor-supplied usage data was for making subscription decisions (94% of respondents), followed by justifying expenditures (86%)."1 Usage statistics are not the only variable used in budgeting decisions, but they certainly play a large role, particularly with the release of new statistics modules from companies such as Scholarly Stats, Ex Libris, and Serials Solutions. These tools have functionality that when combined with the cost of items managed in the ILS offer a look at cost-per-use and other more granular types of data.

Overall, what we can conclude from this informal survey is that the collection of vendor-supplied usage statistics continues to have a significant manual element. While there has been a steady increase in the number of publishers and vendors who are COUNTER-compliant, some content providers have still not adopted this standard. Additionally, the SUSHI protocol is elegant, but has not yet enjoyed widespread adop-

tion. Even those libraries that outsource usage statistic collection often have to manually collect statistics for the "as needed" request. Finally, usage statistics are clearly used for collection development decisions.

# **Interviews with Practitioners**

In an effort to more fully expand on the findings from the survey, the authors conducted three follow-up interviews with librarians at three Association of Research Libraries (ARL) institutions. When asked what was the biggest challenge facing her with regard to usage statistics, one librarian noted that her institution outsources to a third party for the collection of usage statistics, and at times she is uncertain whether or not the statistics are completely accurate. Because she outsources, it allows her to concentrate her energies elsewhere, yet outsourcing inherently means a certain loss of control. For example, publishers are not exempt from making mistakes with their usage statistics. It is not uncommon to see posts to e-mail lists from publishers or content providers indicating that a certain time period had incorrect usage reported and libraries will therefore need to adjust their historical counts. When usage statistics collection is outsourced, it is important to verify that the third party monitors publisher posts such as the one described above and goes back to correct historical counts for its customers.

A challenge mentioned by all three librarians was the amount of manual effort still required for usage statistic collection. Two of the three were using a third party to assist with usage statistic collection and aggregation, but all three saw a manual element remaining for the foreseeable future. There will continue to be those publishers who are not COUNTER-compliant, and there will continue to be those "as needed" requests from administration and from subject specialists. As a segue from the manual component, one librarian questioned the cost-benefit analysis for the amount of time spent on collecting, aggregating, and analyzing usage statistics at her institution. Particularly for "core" electronic resources in a given field—for instance, an electronic resource subscription required for accreditation of some sort—is it truly worth the time and effort to collect monthly statistics?

The time required for any task must be closely evaluated, especially for already overtaxed librarians charged with the acquisition and maintenance of electronic resources. In a survey of electronic resources librarian job position announcements reported in a 2007 issue of Collection Management, Albitz and Shelburne describe the typical e-resources position: "In order to ensure that all ER [electronic resources] activities are accomplished, administrators have had to design positions which only a superhuman could perform successfully, or create 'kitchen sink' positions which include a variety of disjointed responsibilities that no current employees perform."2 Indeed, the management of electronic resources is chaotic and ambiguous, and for the three librarians interviewed, usage statistic collection and analysis was limited to the bare minimum of what was required by administrators, subject specialists, and outside agencies such as the Association for Research Libraries. Echoing the comments received in the survey, two of the interviewees noted that they would like to implement SUSHI, but they currently did not have the time and things were "working okay" as is.

One librarian noted that at her institution, there was a subject specialist in the sciences who is a "usage stat fiend"; this particular librarian collected and analyzed a number of publisher and content provider statistics herself. Perhaps that is one answer to the overworked e-resources librarian—the collection at the larger level as required by outside agencies or administrators is conducted centrally and the remaining tasks, for renewal and cancellation decisions, are delegated to the subject specialists. Of course, this becomes problematic because usage statistics available at vendor websites are often intertwined with other functionality in the administrative modules; thus, distribution of sensitive administrative module usernames and passwords is a delicate matter.

E-books introduce another layer of complexity to the picture. Collecting statistics for e-books is challenging for librarians who are already overtaxed with usage collection for other e-resources, so one librarian who was interviewed mentioned that she is considering patron-driven acquisition. Patron-driven acquisition is a model that has been in existence for a least the past decade but has been slow to catch on. However, it has recently enjoyed more success with the emergence of more e-book aggregators, such as ebrary, EBL, MyiLibrary and others. Sue Polanka generally defines patron-driven acquisition: "Working with a selected vendor, librarians establish a profile based on LC classification, subject, educational level, publication date, cost, and other criteria. E-book titles that match the profile are then shared with the user community via MARC records in the catalog. Once a patron discovers and views an e-book, it is purchased for the collection. There are several iterations of this purchasing model, depending on the selected vendor."<sup>3</sup>

It is important to note that there is a fairly high degree of variability among the models available from the vendors. For example, the vendor may offer customization of the number of e-book patron views before the content is automatically purchased. Regardless of the model, the librarian is still in charge of customizing the profile for his or her institution based on local needs. Patrons contribute to the acquisition process through a controlled environment. With patron-driven acquisition, use literally drives the selecting and purchasing process. In addition to going direct to e-book aggregators, several of the book jobbers, such as Yankee Book Peddler and Coutts, offer patron-driven acquisition. In much the same way that serials agents have had to reinvent themselves with the explosion of e-journals, book jobbers are beginning to see a similar need for reinvention with the continual increase in e-book acquisition.

Not surprisingly, all three librarians explained that usage statistics are one criterion considered at the point of renewal for an electronic resource. Usage statistics are not the *only* consideration, however. Perhaps if the subject librarian is ambivalent about the renewal of a given resource, usage statistics might play a larger role in the decision-making process, but again, usage is rarely, if ever, the only criterion. According to one of the interviewees, extremes raise a red flag, thus causing her to dig deeper, to discover why a resource is enjoying very high usage or experiencing very low usage.

The analysis of usage statistics can be highly localized as well. If an institution has a strong program in a fairly esoteric subject with a very productive but small faculty, then low usage can be justified. Moreover, if one of these productive faculty members is high profile in his or her field, variables that are less easily measured and collected contribute to the renewal decision. The use of the database by one person could attract students to the institution as well as contribute to the awarding of grants and other external funding. The institutional return on investment for the purchase price of the e-resource can be much greater than what is reflected in the usage statistic report.

One librarian who was interviewed mentioned that she is particularly intrigued by the deeper usage analysis made possible by projects such as Project MESUR and variables like the Eigenfactor. She noted that straight

33

usage statistics do not capture the relational aspects among content, so she is encouraged by Project MESUR and Eigenfactor, as well as the introduction of relational recommender tools such as Ex Libris's bX Recommender Service. The wealth of usage data at our fingertips allows researchers to explore the relationships among content, thus paving the way for a broader and more complete usage picture.

## For More Information

Project MESUR www.mesur.org/MESUR.html

*Eigenfactor* www.eigenfactor.org/whyeigenfactor.htm

bX Recommender Service www.exlibrisgroup.com/category/bXOverview

Usage statistics are another item in the storehouse of supplies that the electronic age has given us. Initiatives such as COUNTER have normalized statistics across vendors and publishers and cleared the way for third-party providers to create and sell additional products and services to librarians. While progress has certainly been made in the usage statistic arena, much work is left to be done, such as decreasing the amount of manual effort still remaining and expanding the scope and crosspollination of statistics across varying silos.

# What Is Next?

To determine what is the next evolution in the usage statistic saga would require a visit to the oracle at Delphi. In an effort to survey what is the next trend in usage statistics, Douglas King, in an article published in the *Journal of Electronic Resources Librarianship* in 2009, asked a publisher, several librarians, and Peter Shepherd, the director

of COUNTER, that very question. Their answers reflected what we have discussed in this Library Technology Report. What follows is a very cursory summary of their responses; the full article, and Volume 21, Issue 1 of the Journal of *Electronic Resources Librarianship*, offer a wide-ranging and intelligent discussion of many of the issues surrounding use and usage. Those that King interviewed noted that there will be greater emphasis on the need for use data at a more granular level-the article level-and that will impact how we measure journal quality. As with higher education and other arenas in general, librarians are being asked to assess the value of resources in real and understandable terms, so there will be more emphasis on demonstrating the institutional or community return on investment via all types of library metrics. The economic downturn of 2008 and beyond has meant the analysis of usage statistics leads to very real cancellation and collection decisions. Greater adoption of SUSHI and outsourcing to third parties will significantly reduce the arduous and manual task currently performed in house in many libraries. Perhaps one librarian said it best: "I have no doubt that five years from now, we will look back at the present day and marvel at how much institutional effort was required to do what will be done then, in the near future, nearly effortlessly."4 Indeed, one hopes he is right.

# Notes

- Gayle Baker and Eleanor J. Read, "Vendor-Supplied Usage Data for Electronic Resources: A Survey of Academic Libraries," *Learned Publishing* 21, no. 1 (Jan. 2008): 51.
- R. S. Albitz and W. A. Shelburne, "Marian through the Looking Glass: The Unique Evolution of the Electronic Resources Librarian Position," *Collection Management* 32, no. 1/2 (2007): 16.
- 3. Sue Polanka, "Off the Shelf: Patron-Driven Acquisition," Booklist Online website, www.booklistonline.com/ default.aspx?page=show\_product&pid=3226359 (accessed May 17, 2010),. originally published in Jan. 1, 2009, Booklist.
- 4. Douglas King, "What Is the Next Trend in Usage Statistics in Libraries?" *Journal of Electronic Resources Librarianship* 21, no. 1 (2009): 8–9.