

2013

Analyze This--Usage and Your Collection: Usage Statistics at the Point of Need: Developing a Collaborative Electronic Usage Statistics Program

Anita K. Foster
Illinois State University

Kathleen McEvoy
EBSCO Information Services, k.mcevoy@ebSCO.com

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>

 Part of the [Library and Information Science Commons](#)

Recommended Citation

Foster, Anita K. and McEvoy, Kathleen (2013) "Analyze This--Usage and Your Collection: Usage Statistics at the Point of Need: Developing a Collaborative Electronic Usage Statistics Program," *Against the Grain*: Vol. 25: Iss. 6, Article 42.
DOI: <https://doi.org/10.7771/2380-176X.7437>

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.

Analyze This: Usage and Your Collection — Usage Statistics at the Point of Need

Developing a Collaborative Electronic Resource Usage Statistics Program

by **Anita K. Foster** (Head, Content Acquisitions and Electronic Resources Unit, Milner Library, Illinois State University)

Column Editor: **Kathleen McEvoy** (EBSCO Information Services) <KMcEvoy@ebsco.com>

As library budgets continue to tighten and libraries continue to invest in more on-line resources, it becomes increasingly important to know how electronic resources are being used. It is not sufficient to say a journal or database is useful to support a library's mission. Libraries must provide metrics for the use of such resources to support such a statement. **Milner Library** is no different. The library had avoided major journal and database cuts for many years, but 2013 brought a flat budget and no extra funds to cover inflationary costs or to support new faculty research areas. The library needed to reduce subscriptions to accommodate price increases for core titles. **Milner** began a journal review in Spring 2013 to identify possible journals for cancellation. Unlike previous reviews, librarians had a new metric to add to the tools aiding in making difficult decisions to cancel subscriptions — Project COUNTER statistics. While COUNTER statistics are only one of many types of data available from vendors and just one of many criteria used in the review process, they are the focus of this article.

When the author began working as the Electronic Resources Librarian at **Illinois State University's Milner Library** in 2007, there was no comprehensive program of collecting usage statistics for electronic resources. Database usage was collected and supplied to library administration and the collection development manager, but statistics were not widely disseminated to other librarians. In 2007, availability of Project COUNTER statistics was recent and not all vendors were participating. Within **Milner Library**, some were concerned that not all vendors accurately followed the COUNTER standard for journal usage reporting and therefore accuracy was not as high as it could be. Due to this concern and the time required to find and compile data, journal usage statistics were not collected regularly. It is likely, at the time, that many librarians were not aware of the availability of journal usage statistics, in any form.

Creating a program for collecting usage data can be complicated. This article describes the process developed and still in use at **Milner Library**. Although the current practice uses the UStat statistics software provided by **Ex Libris** to Verde and SFX users, the author believes that most of the steps involved in the process can work in other libraries, regardless of the tools used.



Developing the process at **Milner Library** included the following steps:

- Identify and/or create logins for accessing administrator and usage statistics site(s)
- Evaluate available statistics data. (e.g., is it COUNTER compliant, is SUSHI (Standardized Usage Statistics Harvesting Initiative) harvesting available?)
- Create a master inventory list of vendors/platforms from which to obtain statistics
- Determine a frequency for data collection
- Populate the usage statistics software with data, either from manually loaded files or SUSHI
- Train staff who need access to usage statistics

While the program for gathering statistics was not fully implemented until 2010, the process really began shortly after the author arrived at **Milner Library**. The author quickly discovered that multiple people kept password files for electronic resource subscriptions. Multiple logins were frequently discovered for the same resource. Clearly, the first step in developing the usage statistics program was to unify the diverse sets of administrative credentials. Unifying access under the same login was vital as it reduced the amount of staff confusion when determining if a login was for administrative access, usage statistics access or both. In the beginning, most logins covered both administrative functions (e.g., setting up link resolvers, adding institutional branding) and provided access to usage statistics. As time passed, more vendors provided separate sites for usage statistics, frequently with a new set of login credentials. Initially, user names and passwords were stored in a shared Access database but, eventually, credentials were kept within an electronic resource management system (**Ex Libris' Verde**).

Reviewing available vendor usage statistics was the next step. Understanding what types of data were available is important. Although the COUNTER Code of Practice was first published in 2003 (<http://www.projectcounter.org/about.html>), it was not widely adopted by vendors until after 2006. In 2008, the author began evaluating the usage statistic files, both COUNTER and non-COUNTER, available from various vendors. Many factors were studied, such as timeliness, ease of use, ease of access, and consistent formatting.

Another impetus for looking more closely at journal and database usage was increasing numbers of requests from librarians for statistics of any sort for their subscriptions. **Milner Library** has over 20 subject librarians; it was soon obvious that handling individual requests for statistics could quickly become unmanageable. It was clear that it would be necessary to make statistics available in a single place, easily accessible by all and easy to use. The mechanism for providing statistics was uncertain, however. A couple of usage statistics systems had been identified but had limitations. Providing access to the data in Microsoft Excel or Access files looked like the most feasible method. However, a new system became available which influenced the direction of the project.

In 2009, **Ex Libris** announced the availability of UStat, a Web-based usage statistics system. It was made available to Verde and SFX customers at no charge. UStat utilizes the COUNTER Journal 1 (JR1) and Database 1 (DB1) reports. Excel and text files can be uploaded to the system manually. UStat also has SUSHI (<http://www.niso.org/workrooms/sushi>) capability.

Initially, only the Electronic Resources staff interacted with UStat. The early months were spent learning the system, developing the process for adding data and evaluating the reporting mechanisms available within UStat. One important activity during this step was the development of the list resources for which statistics would be collected. UStat does not have any limitations on the number of platforms included nor on the number of files added to it. With that in mind, the staff developed an Excel file that lists vendors and platforms, what types of COUNTER reports are available, dates when the reports are attainable if they are not available shortly after the start of the next month and any special processing or formatting needs. This file is a constantly evolving document; modifications are made as changes to platforms or usage files occur. It also provides a way to communicate issues about individual platforms between electronic resources staff. See *Figure 1*.

The original plan had been to upload data every month. It was soon clear that the timeline would become unmanageable for a number of reasons. The most significant reason was the time commitment; it could take three to four days to gather and upload data files from the large number of vendors involved. In addition, statistics did not need reporting every month. Quarterly data gathering was determined to be

continued on page 73

Analyze This
from page 72

more appropriate and manageable, and it kept UStat fresh enough to respond to statistics requests at any time of the year.

After the vendor/platform list was created and the process for obtaining files was determined, work began to populate UStat in 2009. The author determined that using 2008 usage data files (when available) as a starting point was logical, as 2008 was the year when a majority of the involved vendors consistently provided COUNTER reports. Within two quarters, the library's UStat account had data in it and was ready for broader release. In June 2010, the author held meetings to introduce the system to subject librarians and to solicit additional information on which vendors to include and to receive feedback on the process as a whole.

Providing direct access to UStat for staff doing collection development was a major goal for the program. As mentioned earlier, more librarians were requesting statistics, but it was increasingly difficult for Electronic Resources staff to provide it in a timely manner. Allowing staff to look up their own information meant the Electronic Resources unit would save time collecting data for individual requests and have more time to spend supporting analysis, if required, and supplying other kinds of statistics, like those available from the library's link resolver. UStat provides a read-only login; having collection development staff use it alleviated any concern about data corruption or loss. Only staff in the Electronic Resources unit has administrative access to UStat.

The final step was training collection development and other interested staff. Although training was intended to demonstrate UStat and its reports, the author also spent time discussing how to analyze the data. A common question was "what does the data {actually} mean?" While such discussions were informative and illuminating, the answer is quite simple. Usage data are just numbers; it is up to the person looking at them to determine the meaning and impact of them within the context in which the statistics are being examined.

The usage statistics program was fully implemented at **Milner Library** in Fall 2010. Throughout the last three years, while the vendor list has been updated frequently due to platform and collection changes, the general process has remained the same. Once SUSHI capability was added, additional vendor and platform data have become available more frequently in UStat.

Now that the program has been in place for three years, what does **Milner Library** know about its subscriptions? Trend information is interesting and one of the first things a user sees after logging in. See *Figure 2*.

continued on page 74

COUNTER							Last Date loaded UStat			
Platform	Resource	JR1	DB1	BK1 of BK2	DB2 available	SUSHI in UStat (JR1 report only)	JR1	DB1	Special Instructions	Note
CSA Burnina									All former CSA Burnina resources are on the new Proquest platform. New usage statistics are retrieved via Proquest.	Login instructions located in ePackage's Admin page
CSWD	CSWD Journals	X					10/7/2013			
EBSCOhost	All available resources	X	X	X		Yes		10/7/2013		
Edinburgh University Press	EUP Journals	X			3rd of next month	Yes			If files are manually downloaded (if SUSHI isn't working), make sure both accounts (PROQUEST and EMLINK) are retrieved when setting reports...	
Electrochemical Society	ECS Digital Library	X	X			Yes				Moving to Highwire Summer 2012. And it has moved. Should this line be deleted? Not at this time. Login instructions are in Verde
Emerald Insight	Emerald Journals	X					10/7/2013			
FirstSearch	All available resources	X	X				10/7/2013	10/7/2013	Delete any record that has a platform for WorldCat Local.	Search for Galegroup
Gale Group	All available resources	X	X	X			10/7/2013	10/7/2013		

Figure 1: List of Vendors/Platforms Used for Gathering Statistics



Figure 2. Initial Dashboard View

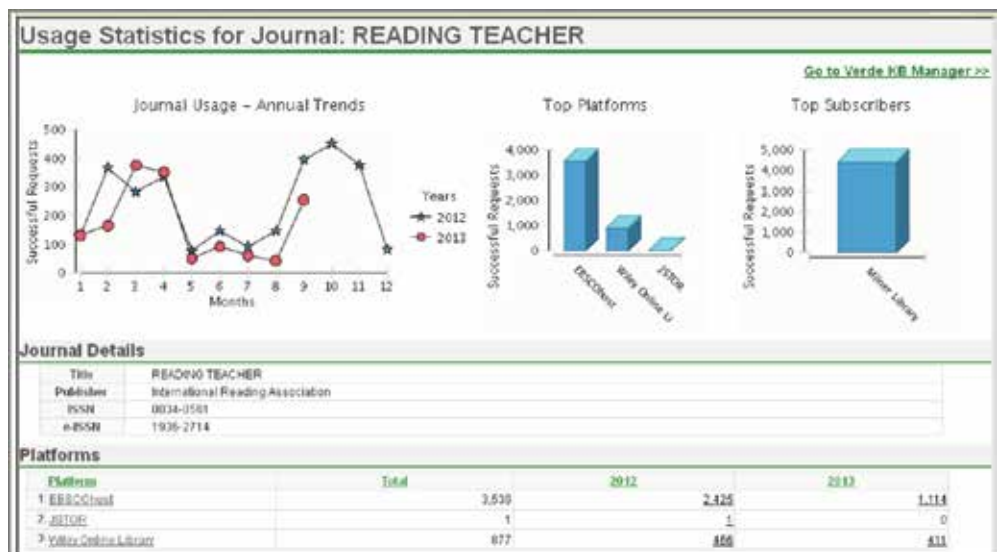


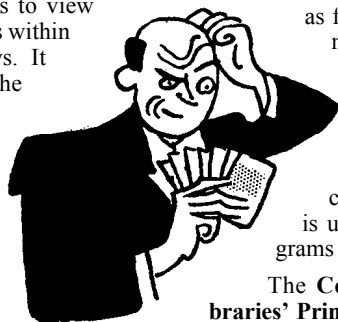
Figure 3. Single Title Usage

It is fascinating to see that the general pattern of usage remains the same, both before and after the introduction of a resource discovery system. Subject librarians have reported that having access to usage data via UStat is very helpful. One librarian used it to present additional information about library support and resources for their subject department's regular program review of its curricula. They also used the data effectively during this year's journal subscription review. One librarian said, "UStat gives us the type of data we need to show faculty how much or how little a journal is used and to provide cost-benefit analysis for specific journals." Following changes in aggregator database coverage of a journal's articles, this same librarian reinstated a subscription to it based on usage data easily accessible to her in UStat.

Many of **Milner's** subject librarians used UStat to support difficult cancellation decisions during this year's journal review. They could easily determine low- and zero-use titles across all of their subscriptions. One of the more useful features is the ability to see use across platforms, when a title is available through multiple resources. This enables librarians to determine if usage in a database is adequate, making a direct subscription less vital. Viewing usage across platforms or resources informs librarians of how and where patrons seek out and use materials. See *Figure 3*.

Although UStat is useful, it is not perfect. At **Milner Library**, demand for usage statistics for streaming media resources is growing, joining the need for better electronic book usage data. UStat does not yet handle either type of resource. Project COUNTER 4, a major revision, goes into effect in December 2013, with reports already appearing in that format. Although report files can be manipulated into a format that will load into UStat, the chances of data corruption increases as more changes are done by a person. If UStat development lags behind **Milner Library's** needs, the library may have to look for a different usage statistics system. However, the process currently in place for gathering statistics should transfer easily as care was taken to create a process not reliant on a specific system.

Electronic resource usage statistics are no longer a mystery for librarians at **Milner Library**. Data is readily available and accessible at the point of need. Using a system like UStat has enabled librarians to view their collections and the materials within those collections in different ways. It is worth repeating though, that the information in UStat is just numbers. Trends can be identified, but it is still up to the individual librarian to assign value and meaning. Nevertheless, it is now easier to determine the value and meaning of those numbers. 🌱



Curating Collective Collections — Data for Collection Assessment at a More Granular Level: ICON As An Example

by **Amy Wood** (Director of Technical Services, Center for Research Libraries) <wood@crl.edu>

Column Editor: **Sam Demas** (College Librarian Emeritus, Carleton College & Principal, Sam Demas Collaborative Consulting) <sdemas03@gmail.com>

In a previous column, **Richard Fyffe** reflected on risks libraries faced as they move toward collective collections or interconnected print collections. What we think we know about the holdings of other collections has grave impact on print retention decisions made locally. As curatorial trends run their course, decisions made without sufficient information can have long lasting, if not permanent, repercussions. The level of detailed information about print holdings required by the community in a rush to clear space is minimal — often title level is sufficient or existing holdings statements in local catalogs disclosing what we think is on the shelf. This article invites us as a community to reflect on how collective collections and other coordinated curatorial efforts can be improved by investing in better data at the risk of slowing short-term gains. It also offers a glimpse at **ICON**, a tool for comparing holdings of newspapers, as an example of emerging best practices in data for collection assessment.

Ithaca set a stake in the ground a few years ago with its **What to Withdraw** tool. The tool uses detailed information about journals at the issue level to support local collection development and management decisions. The foundation is there for applying quantitative methods of rating value, as in the ratio of images and text. A drawback to the tool is the limited data set it supports — **JSTOR** titles from two dark archives that do not allow access to the print.

The **Association of Southeastern Research Libraries' (ASERL)** has developed a noteworthy tool — **Journal Retention and Needs Listing (JRNL)**. **JRNL** was developed for participating institutions to track journal retention commitments between the **Association of Southeastern Research Libraries (ASERL)** and **Florida State University System (SUS)** partners. It is a tool for individual libraries to track their data and a data repository to aggregate a program's data. One drawback is that data is accepted as formatted and is therefore not always consistently expressed. This makes truly automated aggregation of the data impossible. The bigger drawback, for the wider community, is that the tool is unavailable for other programs to use.

The Center for Research Libraries' Print Archives Preservation

Registry (PAPR) takes the realm of tools for aggregating data about print holdings beyond the local. **PAPR** has roughly 50,000 records for 35,000 titles committed for archiving by twelve separate programs (49 combined institutions within those programs). One of **PAPR's** most useful features is the title, holdings and gap reports by program or search results, which users can download. **PAPR** also has a service allowing users to compare a list of publications from their collections with those in **PAPR**. Developing a means to aggregate issue-level data in an automated way is in the works. **PAPR's** big drawback is similar to **JRNL's**, the free-text data fields allow inconsistency of expression of holdings.

As print archiving or shared collection programs mature, **OCLC** continues to improve the tools it offers for collection analysis, and commercial products are being developed as well.

Despite the growing number of tools, the community lacks focus on creating better data. The challenge is getting data in a format which allows us to use it to make better informed decisions. **CRL** addressed that challenge in improving its **ICON** database to better assess newspaper collections. Of utmost concern was providing a tool for the automated comparison of local library collections with the electronic holdings of commercial newspaper databases.

Projects or programs like the United States Newspaper Program,¹ the National Digital Newspaper Program² and the Florida Digital Newspaper Library³ are important examples of how much coordinated efforts can accomplish with regard to collecting and exposing library newspaper holdings. What they lack is information about the holdings of commercial databases and the tools to compare and assess these collections against libraries' print and microform collections. To make decisions about preserving their own collections and purchasing commercial databases, librarians need to know exact holdings down to the issue level and to have at their disposal tools that automate comparisons at that level between collections.

The ICON Database

CRL's primary goals in developing the **ICON** database were:

- to increase the amount and quality of information on newspapers that are and have been published in the U.S. and abroad;
- to increase transparency of commercially produced collections of digital

continued on page 75