



## Electronic Resource Usage Data

Janet K. Chisman

To cite this article: Janet K. Chisman (2008) Electronic Resource Usage Data, The Serials Librarian, 53:4, 79-89, DOI: [10.1300/J123v53n04\\_06](https://doi.org/10.1300/J123v53n04_06)

To link to this article: [http://dx.doi.org/10.1300/J123v53n04\\_06](http://dx.doi.org/10.1300/J123v53n04_06)



Published online: 17 Oct 2008.



Submit your article to this journal [↗](#)



Article views: 100



View related articles [↗](#)



Citing articles: 2 View citing articles [↗](#)

# Electronic Resource Usage Data: Standards and Possibilities

Janet K. Chisman

**ABSTRACT.** Technology is moving rapidly to provide librarians with data needed to make informed decisions regarding the collection management of electronic resources (e-resources). Electronic resource management (ERM) systems provide the ability to store needed basic information about e-resources and associated licensing. The COUNTER protocol provides standardized usage data for journal, database, reference works and books. The SUSHI protocol in turn provides a way to format the data for automated import into ERM systems. Intermediaries, such as ScholarlyStats from MPS Technologies, supply SUSHI-formatted data from an increasing number of platforms via an automated import system. doi:10.1300/J123v53n04\_06 [Article copies available for a fee from The Haworth Document Delivery Service: 1-800-HAWORTH. E-mail address: <docdelivery@haworthpress.com> Website: <<http://www.HaworthPress.com>> © 2008 by The Haworth Press. All rights reserved.]

**KEYWORDS.** SUSHI, COUNTER, ScholarlyStats, usage data, electronic serials management

What to buy and how to justify a purchase is an ongoing challenge for collection development librarians. Serials offer the additional dimension of commitment of money over a period of time. The sums involved are often substantial, ranging from a few hundred dollars per subscription to over a million dollars for some consortial or package deals. Per-title usage information is one essential component that assists in these

---

Janet K. Chisman is Head, Serials and Electronic Resources, Washington State University, Libraries, Serials and Electronic Resources Unit, Pullman, WA 99164 (E-mail: [chisman@wsu.edu](mailto:chisman@wsu.edu)).

The Serials Librarian, Vol. 53(4) 2008  
Available online at <http://ser.haworthpress.com>  
© 2008 by The Haworth Press. All rights reserved.  
doi:10.1300/J123v53n04\_06

purchase decisions. The electronic resource environment provides the perfect venue to supply such usage statistics to help librarians build collections that are cost-effective and focused on user needs. For example, as print indexes and abstracts moved to online access, librarians were given large amounts of usage data, either in paper or electronic form. However, the definitions of the data elements being provided appeared to be unique to each vendor, and the available data elements also varied. For instance, was a “search” at Database A comparable to a “search” on Database B? What exactly was a “search?” With varying data elements and varying definitions, there was no way to compare the use between and among database providers to decide which were providing the access that patrons needed and budgets could afford. The explosion of full-text electronic journals with even more vendors and more variations on the theme of “use,” “search,” “retrieval,” “turnaway,” etc., further complicated the ability to use the data to make informed decisions.

### COUNTER

COUNTER (Counting Online Usage of Networked Electronic Resources) was initiated in early 2002 to make the recording and reporting of electronic usage statistics comply to standard definitions and reporting formats.<sup>1</sup> COUNTER emphasizes the provision of basic usage data rather than the development of increasingly detailed usage reports. The standardization of this information benefits both librarians and vendors. Librarians analyze the data to study patron use patterns, while vendors benefit from adding value to their products by supplying a standard set of data to subscribers.

The first published COUNTER code, available in January 2003, covered journal and database reports. In March 2006 the COUNTER Code of Practice for Books and Reference Works was released. Reports now available include:

*JR1* = Journal Report 1: Number of Successful Full-Text Article Requests by Month and Journal

*JR1a* = Number of Successful Full-Text Article Requests from an Archive by Month and Journal

*JR2* = Journal Report 2: Turnaways by Month and Journal

*DB1* = Database Report 1: Total Searches and Sessions by Month and Database

*DB2* = Database Report 2: Turnaways by Month and Database

*DB3* = Database Report 3: Total Searches and Sessions by Month and Service

*JR3* = Number of Successful Item Requests and Turnaways by Month, Journal and Page Type

*JR4* = Total Searches Run by Month and Service

*BR1* = Book Report 1: Number of Successful Title Requests by Month and Title

*BR2* = Book Report 2: Number of Successful Section Requests by Month and Title

*BR3* = Turnaways by Month and Title

*BR4* = Turnaways by Month and Service

*BR5* = Total Searches and Sessions by Month and Title

*BR6* = Total Searches and Sessions by Month and Service

Not all vendors supply all reports, but JR1 reports are almost universal among compliant vendors.

One indication of the success of COUNTER is the large number of compliant vendors listed on the COUNTER web site ([www.projectcounter.org](http://www.projectcounter.org)) along with the types of reports each supplies. The standard terminology allows usage comparison across vendor reports and provides librarians with usable data to assist in collection development decisions such as deciding to use alternative information delivery mechanisms—online document delivery, for example—for low-use titles.<sup>2,3</sup>

While COUNTER eased the problem of supplying consistent data across vendor platforms, a tremendous amount of staff time was needed to contact each vendor individually and retrieve the data. In addition there were the issues of data storage and manipulation. Many libraries developed in-house spreadsheets and databases specifically to house usage data and often to add more information about a title such as Impact Factor, interlibrary loans, cost, and so forth. These staffing and storage issues highlighted the need for emerging commercial electronic resource management (ERM) systems to automatically import and store the data. The key words here are “automatically” and “store.” Automatic import would save large amounts of staff time and storing would obviate the need for a separate system for this function.

## *SUSHI*

The drive for the automated import of usage data into ERM systems revealed the need for additional standardization not covered by the

COUNTER initiative. While data in COUNTER are presented as clearly defined elements in a series of standard reports, the format of the data varies from vendor to vendor. This variation prohibits automated import since automation requires standardization. Another missing element is the functionality which allows one computer to request data from another computer and for the second computer to respond to the request and supply the data (a Web Service, query and response component).

As individuals worked with the COUNTER data and attempted to import them into existing ERM systems, the limitations of COUNTER became increasingly evident. An informal group of librarians, content providers and integrated library system (ILS) vendors formed to find a solution. The group began discussions in the summer of 2005, was recognized by NISO in November, and by late 2005 had successfully tested the concept of automated import of usage data. The group named the new protocol Standardized Usage Statistics Harvesting Initiative, or SUSHI.<sup>4</sup> The SUSHI protocol is in a trial period until May 20, 2007.

SUSHI is a fairly simple protocol using existing web-based products (Web Services, SOAP or Simple Object Access Protocol, and XML schema) to request, format, and import usage data. The following brief overview of the SUSHI import process is based on NISO Webinars that contain background information for librarians and vendors as well as detailed examples of the XML request and response.<sup>5,6</sup>

The import process begins with a library with an ERM system plus associated SUSHI client software paired with a content provider with SUSHI server software and access to the COUNTER usage data. The ERM system sends a request to the SUSHI client—often housed on the same computer—for a COUNTER report. The SUSHI client prepares an XML request and sends it via the Internet to the content provider's SUSHI server, which reads the request, processes the requested COUNTER report into XML format, and prepares a response message formatted according to the SUSHI XML schema. The COUNTER report in XML format is added to the response and both are sent via the Internet to the SUSHI client at the requesting library. The SUSHI client processes the response and extracts the data. The data are then passed to the ERM system and processed for use/display.

At this writing, vendors participating in SUSHI include content providers with SUSHI servers that supply data: EBSCO Information Services, Project Euclid, Swets Information Services, and Thomson Scientific; and ILS vendors with SUSHI clients that receive data: Ex Libris and Innovative Interfaces, Inc (III). III has announced successful SUSHI loads into their ERM system. Other vendors actively working in this

area include Ex Libris, Otto Harrassowitz, OCLC, Serials Solutions, and SirsiDynix.<sup>7</sup>

For any new standard to be successful, it must be accepted and used by the parties involved. Therefore, the next step in the successful implementation of SUSHI is a wider adoption by content providers and ILS vendors. Librarians have a prime role here in raising awareness of SUSHI, emphasizing the need for this initiative and encouraging the adoption of the protocol by noting the competitive edge companies will gain by early implementation. This librarian advocacy will help move SUSHI into the same position now held by COUNTER—something fairly standard that content providers and ILS vendors implement because it is good for business. Information on what is required and how to implement SUSHI can be found at the SUSHI web site ([www.niso.org/committees/SUSHI/SUSHI\\_comm.html](http://www.niso.org/committees/SUSHI/SUSHI_comm.html)).

### *ScholarlyStats*

As librarians promote wider implementation of SUSHI, a new company, ScholarlyStats from MPS Technologies, has stepped in to fill the need for SUSHI-formatted data as well as to supply other value-added services.<sup>8</sup> ScholarlyStats is working with content providers to set up agreements that allow them to gather usage statistics from the content providers for ScholarlyStats' subscribers. ScholarlyStats collects the usage data reports, standardizes them, and puts data through additional cleanup processes to enhance the quality of the information. ScholarlyStats works with ILS integration partners to develop automated import of the data into their electronic resource management systems. To date, III's ERM system has successfully imported data from ScholarlyStats. The company is currently in beta test with Harrassowitz' ERM system, while Ex Libris is in development work with its ERM system, Verde, to support the data transfer. Another integration partner is Thomson Scientific. ScholarlyStats and Thomson Scientific have successfully integrated Thomson's journal use data with the usage data from ScholarlyStats. In June 2006 ScholarlyStats and Thomson Scientific announced a successful import of usage data for the University of Melbourne from 21 vendors into Thomson Scientific's Journal Use Reports (JUR) system using the XML SUSHI protocol.<sup>9</sup> Several libraries are now using this functionality. This extends the information available for collection analysis by coupling usage data with Impact Factor and publication information from Thomson Scientific's Journal Use Reports.

ScholarlyStats provides additional services beyond monthly collection and automated dissemination of journal and usage statistics. They administer all usage statistics platform access details in one place, perform additional data cleanup, provide consolidated reports across platforms, offer a suite of custom reports that summarize usage, and allow access to usage statistics by library staff.

### ***PUTTING IT ALL TOGETHER— ONE LIBRARY'S EXPERIENCE***

Washington State University (WSU) is a mid-sized land grant university located in the agricultural region of eastern Washington. The WSU Libraries followed a fairly common path in dealing with data associated with electronic resources.<sup>10</sup> Librarians were frustrated in the mid-1990s trying to interpret usage data supplied by vendors where the definitions varied widely and there was no ability to compare statistics. We developed a database to assist in making serials selection/deselection decisions. The maintenance of this system was very labor intensive, and frequent upgrades in software and hardware caused many access problems.

Budgets were tight, and more informed decisions were needed. We hungered for a more organized way to manage the data and were fortunate in 2002 in joining the Innovative Interface ERM module development team. When the product became available, we purchased it and began our implementation in earnest with reorganization in October 2004 and the formation of a Serials and Electronic Resources unit within Collections and Technical Services (CTS). While we appreciated the ability to store basic data about our electronic resources and licenses, we were looking forward to being able to generate a cost per use within our integrated system. So we welcomed the opportunity to beta test the automated import of SUSHI usage data from Project Euclid in April 2006. Beta testing revealed that the Customer IDs assigned by Project Euclid were too long for the III ERM configuration area. This problem was quickly solved by III changing its system to accommodate a longer Customer ID. In beta testing III began by running the harvests manually first and when this worked, they moved to testing an automated load. It only took three days to get the automated load to run also. This automated load showed us the potential, but we still needed SUSHI-formatted data from most of our vendors.

Thus, when we heard of a new company—ScholarlyStats, supplying data in SUSHI format—we realized it was the right company at the right

time. Since we had just participated in the beta test of the automated import of SUSHI-formatted data into our ERM, the timing of the announcement was fortuitous. A quick e-mail to III resulted in a three-way partnership. III and ScholarlyStats would test importing to the III ERM module using the WSU system. III is also developing the capability to tie the usage data imported into ERM with the payment information so we can generate a cost per use from within our integrated catalog. Work with III and ScholarlyStats began in August 2006. The first test downloads were completed in October. More details needed to be resolved to allow the download of data from more than one vendor on the same day. This was accomplished in January 2007. In February, data from 34 different vendors (representing over 150 publishers and 35,000 titles) for the previous 12 months are scheduled for download into our ERM.

A considerable amount of system work was required to accomplish this import of statistics from ScholarlyStats into the III ERM system. The exchange of information for testing was slowed by the location of the ScholarlyStats programmer in India, limiting the exchange of emails to one per day as the programmers moved through the troubleshooting process. ScholarlyStats had to develop a method to include our account number and the platform name in the customer ID field. This allowed us to create a separate entry for each platform in our configuration field and for the imported data to locate the correct access provider for the harvested data. The first test revealed that the ScholarlyStats SUSHI server expected a Requestor ID (distinct from the Customer ID) to allow the SUSHI server to determine if the request was being made "on behalf of" the Customer ID. WSU, however, was making the request directly and not using a third party or ASP service, and III was using SUSHI 0.1 with the Customer ID and the Requestor ID values the same. When III implements SUSHI 1.0 they intend to make the Requestor ID/Customer ID configuration more flexible. Both companies had to make changes to their systems, and it took about seven weeks before we had our first successful harvest of usage data for Elsevier ScienceDirect titles.

We then tested each platform on a different day and worked through additional problems that arose. Failures of several loads were caused by the time difference between the settings for the downloads on the ScholarlyStats server and those on the WSU client. Another problem that was identified during beta testing was the inability of ScholarlyStats to handle more than one request from WSU per day due to the need to make manual changes in the system to change the Requestor ID each time a request was made from the WSU III ERM. The solution for this was found, and WSU Libraries moved into the ScholarlyStats production

environment in late December. In January 2007 we were able to import statistics for all of 2006. In February 2007 we encountered another download problem which was quickly resolved.

The work in setting up the needed configuration table in III ERM and entering information into the ScholarlyStats system took several days but was minimal compared to the repeated manual downloading of this information from 34 different vendors and subsequent integration into a spreadsheet. For ScholarlyStats, we provide our customer id and password for each subscribed platform along with the URL to the usage statistics web site. Special instructions can also be included as needed. We estimate a .5 FTE position is now available for other duties because of this change in procedures.

On the ERM system side, information needed to make the request at a specified time is entered into a configuration table. This information includes a code for the content provider which is used to link the incoming data with the correct resource record(s) in the system, a customer id, day of the month to make the request, number of months data to request, and a Web Services Description Language (WSDL) URL—that lets the SUSHI server at ScholarlyStats know what services we are requesting.

Currently the end product of all this beta testing and setup is the ability to retrieve a resource record for a specific data supplier, such as Elsevier ScienceDirect. Opening the Usage Statistics option in the record delivers a screen showing the months of data available. An Export link opens an Excel spreadsheet with a summary of the payments for the year, total uses and cost per use for the year to date at the top. Then each title is listed with its own calculated cost per use, uses for each month covered, and total uses for the entire time period covered. This automated retrieval and calculation of cost per use is invaluable in the efficient management of usage data.

Cost per use is just one factor considered in the selection/deselection process, however. The latest iteration of our serials database is an Excel spreadsheet with the following elements being tracked: Title, ISSN, III Record Number, Fund Code, Cost, JCR Total Cites, JCR Impact Factor, JCR Immediacy Index, JCR Articles, JCR cited Half Life, and Data Usage Statistics for each vendor. A cost per use is generated as needed for each title based on the subscription fees for the year and the usage statistics imported from each vendor. The title, ISSN, fund codes and subscription costs (by limiting to a paid date range) are pulled from the order records via a create list functionality from the Acquisitions system. This information is imported into the Excel spreadsheet using a series of macros that format the information as needed. The usage

statistics are imported annually or semiannually from each vendor supplying COUNTER-compliant statistics. The data are further massaged to produce a usable spreadsheet.

A new serials decision database is currently being developed by our new Head, Collections and Acquisitions with a greater WSU focus and should be operational in April 2007. We are dropping the generic JCR data from the database and including instead the number of WSU-authored articles published in a journal, the number of times articles in a journal title are cited in a WSU-authored article, and the number of interlibrary loan photocopy requests are placed by WSU faculty, staff and students for articles from a journal title. Access to the dropped JCR information is available by linking to the Ulrich's Serials Analysis System. Maintenance is still labor intensive but the increased focus on data authored and requested by WSU faculty, staff and students will improve our ability to identify and subscribe to publications more relevant to faculty and student research and teaching needs.

The ideal situation would be to have the ability to store the type of information in the new serials decision database in the existing ERM module or a new collection management module of the integrated catalog. We are moving in that direction with the ability to import usage statistics and associate them with costs with the paid date(s) for the same year. Unfortunately, we usually pay a subscription in the year before the subscription actually begins, so our cost data are currently a year off. That is, a subscription beginning in 2007 is paid in 2006, so the payment costs are associated with the usage data for 2006 instead of 2007. III will be enhancing the ERM with a new field in the order record to record the subscription period. This will then associate the correct payment with the imported statistics. If we could then enhance the ERM system further to provide storage of additional custom data (for us it would be the WSU-authored and cited articles and number of ILL photocopy requests by WSU affiliated patrons) as defined by each individual library, we would have the beginnings of a very usable collection management system within the integrated catalog. We could then eliminate the stand alone serials decision database and use our integrated catalog as the central store of data associated with e-resource collection management.

### **CONCLUSION**

Technology is moving rapidly to provide librarians with data needed to make informed decisions regarding the collection management of

electronic resources. The development of ERM systems is providing the ability to store needed basic information about e-resource suppliers and the associated licenses. The COUNTER protocol provides for standard usage data for journals, databases, reference works, and books. The SUSHI protocol then provides a way to format the data for automated import into ERM systems. Currently there are a limited number of vendors supplying SUSHI-formatted data directly. However, librarian action will ensure the widespread adoption of the SUSHI protocol by content providers and ILS vendors. In the interim those with large collections or a pressing need for usage data in SUSHI format may consider use of an intermediary, such as ScholarlyStats, that can supply SUSHI-formatted data from an increasing number of platforms via an automated import system.

Vendors who have produced ERM systems are developing the means to integrate these imported data with subscription cost information already available in the acquisitions modules of the ILS. This will make available for one more vital piece of information—the cost per use for electronic resources. Integration of this information with journal citation data provided by other third parties will expand even further the ability to make informed decisions regarding collection development.

My intent here has not been to promote a specific product, but to emphasize how the development of these products, in association with the COUNTER and SUSHI protocols, is bringing us closer to the goal of successfully using our usage data reports to improve and focus our collections. I also look forward to a standard license with a set of standard portable descriptive data elements that librarians can automatically load into ERM systems. I propose we name it Absolutely Terrific License Attributes Simply Transferred—AT LAST.

## NOTES

1. "About COUNTER," COUNTER, <http://www.projectcounter.org/about.html>.
2. Oliver Pesch, "Ensuring Consistent Usage Statistics, Part 1: Project COUNTER," *The Serials Librarian*, 50, no. 1/2 (2006):147-161. doi:10.1300/J123v50n01\_14.
3. Alfred Kraemer, "Ensuring Consistent Usage Statistics, Part 2: Working with Use Data for Electronic Journals," *The Serials Librarian* 50, no. 1/2 (2006):163-172. doi:10.1300/J123v50n01\_15.
4. NISO National Standards Information Organization, "NISO Standardized Usage Statistics Harvesting Initiative (SUSHI)," NISO, [http://www.niso.org/committees/SUSHI/SUSHI\\_comm.html](http://www.niso.org/committees/SUSHI/SUSHI_comm.html).

5. Adam Chandler, Phil Davis and Tim Jewell, "Introduction to SUSHI for Librarians and Content Providers," *NISO Webinar*, (May 17, 2006), <http://www.niso.org/committees/SUSHI/SUSHIWebinarslides051506.pdf>.

6. Ted Fons and Oliver Pesch, "SUSHI: The Technology Unveiled," *NISO Webinar*, (May 17, 2006), <http://www.niso.org/committees/SUSHI/SUSHIWebinar052406.pdf>.

7. Adam Chandler and Tim Jewell, "The Standardized Usage Statistics Harvesting Initiative (SUSHI)," *Serials*, 19 (March 2006):68-70.

8. "Simplifying Statistics," ScholarlyStats from MPS Technologies, <http://www.scholarlystats.com>.

9. "ScholarlyStats and Thomson Scientific Announce Successful SUSHI Test for University of Melbourne," NISO, [http://www.niso.org/committees/SUSHI/ScholarlyStats\\_ThomsonScientifi.pdf](http://www.niso.org/committees/SUSHI/ScholarlyStats_ThomsonScientifi.pdf).

10. Janet Chisman and John Webb, "Electronic Resource Management Using a Vendor Product," *Journal of Electronic Resource Description & Access*, 8, (forthcoming).

doi:10.1300/J123v53n04\_06

