

This article was downloaded by: [University of Oregon]

On: 28 October 2009

Access details: Access Details: [subscription number 906611115]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Journal of Electronic Resources Librarianship

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t792303957>

### What is the Next Trend in Usage Statistics in Libraries?

Douglas King

Online Publication Date: 01 January 2009

**To cite this Article** King, Douglas(2009)'What is the Next Trend in Usage Statistics in Libraries?', Journal of Electronic Resources Librarianship,21:1,4 — 14

**To link to this Article:** DOI: 10.1080/19411260902858276

**URL:** <http://dx.doi.org/10.1080/19411260902858276>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## E-OPINIONS FROM THE FIELD

---

### WHAT IS THE NEXT TREND IN USAGE STATISTICS IN LIBRARIES?

---

**Douglas King**

In answering the question “What is the next trend in usage statistics in libraries?” an eclectic group of respondents has presented an assortment of possibilities, suggestions, complaints and, of course, questions of their own. Undoubtedly, usage statistics collection, interpretation, and application are areas of growth and increasing complexity in our field, rife with opportunities for advancement and improvement. The five contributors approach the topic from varying angles and give this reader—and hopefully others—plenty to consider. I am grateful to the writers for their thoughtful and creative contributions.

***MARTHYN BORGHUIS, SENIOR MANAGER, ACADEMIC  
AND GOVERNMENT PRODUCTS, ELSEVIER***

The production and distribution of usage data to libraries has been a great success. Some twenty years ago when I worked as a head of the social sciences library at State University Groningen (Netherlands), lending figures and reshelving studies were the most commonly used techniques to get at least some quantitative insight into the usefulness of the collections maintained. Collection development at the time was often driven by a library committee consisting of research staff, students, and active library visitors. At the change to a new millennium, publishers started offering e-journal and e-book collections to universities, research centers, hospitals, and corporations. Since 2001, many publishers have decided to make usage data of these collections available to subscribing institutes. I remember very clearly when the first simple usage reports were dispatched by e-mail as an Excel attachment.

In the years thereafter, a wealth of e-journal usage data became available. The provision of usage reports by publishers to individual institutes strongly improved the local collection evaluation process. Since that time, librarians do understand much better and in greater detail their

institutions' information need. As a result, publishers benefit strongly from this development: Readership levels of journals influence the editorial tasks considerably, and journal management staff takes advantage of the wealth of usage data.

In 2002, Project COUNTER (<http://www.projectcounter.org>) was founded, providing standards for the format and content of usage reports on the level of individual journals and, since recently, also of books. Detailed requirements for report format, layout, and timeliness of usage reports contribute strongly to the collection evaluation process. Features such as reporting zero usage and the article double-click filter have increased the reliability and usefulness of the COUNTER standards maintained. Librarians and publishers now share the same usage data, which makes the license negotiation process a lot more fact-based.

For the next couple of years, it is expected that new usage metrics will be developed, based on the usage per article, and thus will add new criteria for what may be called the attractiveness, or usage intensity, of a journal or book. High on the COUNTER to-do list is the development of a COUNTER Usage Factor, an equation of full text article usage published in a journal during a particular period and divided by the number of articles available online during that same period. Similarities with the Impact Factor per journal, produced by the Institute of Scientific Information, are of course obvious, but the working group managing the project expects quite different results. This COUNTER project is in a final stage of preparation. A selection of publishers will be asked to provide the necessary article usage and article reference data.

Research into the optimal period to cover is the key part of the Usage Factor research project. Another topic will be to define which article types (and their usage) should be included or excluded from the numerator and denominator of the Usage Factor equation. Where the Impact Factor normally runs up to two years behind (it needs the citation data of articles that were published in the two years before), the COUNTER Usage Factor will be more up to date as usage data can be collected immediately after publication. It is the definition of the period for which both number of articles available online and the usage data of these articles will be collected, so quite some experimentation will be needed before this metric will go live.

***PAUL BRACKE, ASSOCIATE DEAN FOR INFORMATION  
TECHNOLOGY AND RESOURCE SERVICES, PURDUE  
UNIVERSITY***

Academic libraries face a challenging environment that makes the collection and analysis of usage statistics more important than ever. Colleges

and universities have many competing demands for funding but are not seeing increases in funding, particularly in the current economic climate. This makes demonstrating value and impact to the campus essential and is demonstrating that the resources we do have are being used effectively and efficiently. There are also exciting opportunities for libraries to make use of usage data to design new services for our users. Although there are many ways in which library practice might become more data driven and evidence-based, I would like to highlight two trends in library statistics that will be particularly important.

The first point is that libraries must do a much better job of demonstrating their value and impact to their campuses. There has been some movement in the direction of measuring impact, and not merely inputs and outputs. User-centered, quality-focused measures such as LibQUAL+ and other assessment tools emerging from the Association of Research Libraries (ARL) New Measures Initiatives, for example, demonstrate outcomes that go well beyond volume counts. Statistical measurement and analysis of library programs need to go farther, however, and evaluate the effectiveness of library programs in a campus context. Libraries must also be able to show how their programs have an impact on campus goals, such as learning outcomes, student persistence, and research income. Although it does not employ statistical analysis, a recent University of Illinois study, for example, is illustrative of this conceptual approach. The study demonstrates the value of libraries by calculating a return on investment in terms of sponsored research funding generated. Although the specific measures may vary from campus to campus, libraries will need to develop statistical analyses that quantify impact on the entire campus. The development of new measurement tools and models that do this will be one thing to look for in coming years.

The future of library statistics is not all administrative, however. There are also some opportunities to create new, or enhance existing, services using data. To some extent, libraries have been doing this for years, but there is much more that can be done. Most libraries make use of circulation or access statistics to make collection decisions, usage statistics may be used to make decisions about facilities or staffing levels at service points, and Web logs may be used to improve Web site design. These examples require intervention by staff, however, and have an impact at a broad level—the collection or library or Web site changes for all users based on aggregate usage. “My Library” sites allow for a greater amount of immediacy in providing users with their preferred interface, although much of the customization is selection either by the librarian or a user. Though this is an improvement, users are growing accustomed to interfaces that add value to their experience through analysis of usage data—features such as personalized recommendations or “More Like This” functions, features that

require only that a user make use of the site. Libraries have been reluctant to pursue these opportunities, but analysis of aggregate data alone can result in service improvements. The University of Huddersfield, for example, has developed a recommendation system by mining circulation records and has even released the data for download. Recommender systems and other tools for enhancing access to library resources and services, even down to the truly personalized level, could be built using statistical techniques to analyze usage patterns. Although the privacy issues involved are not trivial, making usage statistics work to enhance access for users will be of increasing importance.

**DAVID FOWLER, HEAD OF ACQUISITIONS, UNIVERSITY OF OREGON**

The current economic downturn, though unexpected in its severity, is beginning to show predictable consequences in public universities and, by extension, university libraries. The State of Oregon, the Oregon University System, and the University of Oregon are no exceptions. Although to date we have fared somewhat better than our academic brethren to the north (in Washington) and to the south (in California), we have been impacted significantly by ongoing negative economic processes over the course of several years, which has been exacerbated by the loss of confidence in financial markets and the subsequent tightening of purse strings at the state level. The current downturn notwithstanding, over the past several years, serials inflation has continued apace at about 8.8% annually; our materials budget has remained relatively flat, with only occasional one-time budget augments being dispersed, which have only delayed the inevitable; and, to date, there has been no progress in creating a new university budget model that factors serials inflation costs into our budget on an ongoing basis.

The unavoidable impact of all this is going to be the necessity of canceling serials to balance our budget. There has not been a focused serials cancellation effort at this institution in several years, so this also means that this will be the first serials cancellation project that necessarily and unavoidably will seek to evaluate and identify what are now primarily electronic serials, including both e-journals and databases, for potential cancellation.

This leaves the library in the position of seriously using electronic statistics for the first time as tools to make quantitative and somewhat qualitative evaluations but with only moderate and uneven institutional expertise in massaging these statistics into useful patterns and formats for the purposes of these evaluations. Although the final decisions will, of course, be made by professional librarians in consultation with the

teaching and research faculty, this library essentially made the decision a year ago to attempt to begin outsourcing our statistics-gathering work to Serials Solutions in the form of the new 360 Counter tool. Though initial reviews of the out-of-the-box quality of this particular tool in the library world have been mixed, there is little doubt that Serials Solutions and other companies will continue to develop and improve this and similar interfaces and services.

The aforementioned budget issues, combined with a potential lack of adequate staffing or staff expertise, I would argue, indicate that like many other aspects of business life, library statistics gathering will continue to be outsourced and that libraries will have more and expanded statistical services being performed on their behalf by outside companies, whether Serials Solutions or another library services company.

Currently, owing to a lack of critical mass of SUSHI-compliant vendors, many of these statistics are still hand-loaded into these tools by library staff members. As a Standardized Usage Statistics Harvesting Initiative (SUSHI) becomes the prevalent standard in the industry, it seems that most vendors of any size will have adopted it within a fairly short amount of time. The bottom line is that the process of statistics gathering, currently requiring library staff to manually go into vendor Web sites at periodic intervals to pull statistics of varying quality and standardization (sometimes at the prompting of vendor e-mails), can soon move to a near-completely automated process, whereby a tool such as 360 Counter will automatically query vendors and compile statistics.

I anticipate that quite soon, once a library has registered its format preferences with the vendor, the entire process, including manipulation of data into various desired formats, will be accomplished in a completely “hands-off” fashion and completely “offshore” from the library’s walls. We will then only have to pull an awaiting and up-to-the-minute report whenever we need various iterations of cost, usage, and cost-per-use data to make our collection development data. I can even see the potential of such data someday soon being transmitted instantaneously to these statistical tools, upgrading the currency of the data to as-yet unattainable heights.

All of this is already beginning to happen, and I fully expect this technology to mature very quickly. Thus, I expect that libraries themselves will likely be happily and permanently out of the statistics-gathering business themselves within five years, which will then free up staff labor and time to be concentrated in other vital areas of the library acquisitions and collection development operations.

The recent state of statistics gathering has been fairly arduous but, nonetheless, it has been far easier, and far more standardized, than it was even five years ago. 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.

**JAN LEWIS, ASSOCIATE DIRECTOR, EAST CAROLINA  
UNIVERSITY**

Benchmarks are important to librarians. We like to see how our library's collections, circulation figures, and user satisfaction compare with those of our peers. We use WorldCat Collection Analysis, and LibQual+, respectively, to compare our collections and user satisfaction levels. Many of our regional and national associations collect and disseminate annual statistics that allow comparisons of acquisitions and operating budgets, circulation figures, and gate counts. However, when it comes to meaningful comparisons of our database and e-journal use statistics and—perhaps more important—the cost-per-use of these resources, we are left in the cold. A mechanism for collecting and disseminating these data should be the next trend in statistics for libraries.

It could not come at a better time. With the country officially in a recession, nearly all libraries, regardless of whether they are funded publicly or privately, have been affected by falling tax revenues and reductions in the value of their or their parent institutions' endowments. Most are facing declining budgets, with the threat of deeper cuts looming. In this environment, it is critical that librarians have the data needed to assess the value of their online resources and make informed decisions based on a cost-benefit analysis. The phrase "data-driven decision making" may be in danger of becoming an overused cliché in library circles, but it will, in fact, be of even more importance as libraries make tough decisions about which resources to keep and which to cancel.

The ability to compare cost-per-use across a defined set of institutions will aid in a library's assessment of its subscription resources, adding one more element to those that are currently considered, such as the size and level of relevant programs and audiences, curricular and research needs, and accreditation requirements. In the case of e-journals, comparative cost-per-use data may be helpful in deciding whether to switch to a pay-per-use model or provide interlibrary loan or document delivery for a particular title instead of subscribing to it.

An added benefit may be the identification of best practices in e-resources promotion and marketing. A library that has consistently lower cost-per-use figures than its peers may be using more effective marketing methods to reach faculty and students, have a well-designed Web site and effective user-centered information control tools, and/or be successful at integrating library resources into the institution's course management system.

Sharing cost-per-use data will face some hurdles. First, of course, is the need to ensure that data are being calculated in a consistent manner across resources and institutions. It seems logical to look to existing COUNTER reports as a starting point. According to the COUNTER Web site (<http://www.projectcounter.org/about.html>), COUNTER

brings the following benefits to librarians, publishers and intermediaries: Librarians are able to compare usage statistics from different vendors; derive useful metrics such as cost-per-use; make better-informed purchasing decisions; plan infrastructure more effectively. Publishers and intermediaries are able to: provide data to customers in a format they want; compare the relative usage of different delivery channels; aggregate data for customers using multiple delivery channels; learn more about genuine usage patterns.

One could argue that publishers and intermediaries are best situated to calculate cost-per-use. They have access to all the data needed—what the library pays, the resources to which it has access, and usage numbers. Columns could be added to appropriate COUNTER reports that show the cost per search or cost per downloaded article or other item. Then, librarians could be assured that the cost-per-use data was compiled using the same methodology and could compare other libraries' data with their own. Of course, libraries currently can use COUNTER reports to “derive” cost-per-use information, and many do, either using spreadsheets or more sophisticated home-grown procedures or, increasingly, by using commercial electronic resource management systems. However, why not streamline the process by asking publishers and intermediaries to include this information in COUNTER reports?

Elsevier might take the lead in this area. It has recognized the pressure on libraries to show the value they provide and, more precisely, the return on investment of library expenditures on scholarly resources. Elsevier devoted the January 2008 issue of *Library Connect* to information valuation and recently published a white paper entitled *University Investment in the Library: What's the Return? A Case Study at the University of Illinois at Urbana-Champaign*. The study looks at the contributions one academic library made to the generation of grant funding at its university. By adding cost-per-use data to the excellent Web-based usage reports it currently provides, along with a mechanism for subscribers to share the information with other libraries, Elsevier can show its leadership in this area.

Other options for sharing cost-per-use data across institutions include coordination by the ARL or regional library consortia such as ASERL. Or, how about a service modeled on LibQual+, in which libraries that subscribe



to the service during a specific period have access to other libraries' results similar to the way in which libraries that participate in LibQual in a specific period can see other libraries' results notebooks?

Whatever model is used for sharing these data, it will take lots of thought and collaboration to design and implement a workable system. However, in an environment focused on outcomes-based assessment, with limited funds and with an ever-increasing array of electronic resources competing for those funds, it will be worth the effort.

### ***PETER SHEPHERD, DIRECTOR, COUNTER***

The last few years have seen not only great improvements in the reliability of usage statistics for librarians but the emergence of usage statistics as a central component of the librarian's management toolkit. However, much remains to be done. Technology continues its relentless march, and the way the usage of online publications is recorded and reported must take this into account; what was a meaningful metric five years ago may no longer be meaningful in the current technology environment. The main purpose of most librarians is not, we must remind ourselves, the collection and management of usage statistics; the time they can devote to this exercise is limited, and the process should be made as efficient as possible. The current global economic downturn will, inevitably, increase the pressure on library budgets in both the public and the private sector, and tough decisions will have to be made on collections in the next few years; usage statistics should help ensure that these decisions are well informed and evidence-based. Finally, the technical and business models for online publishing, especially of journals, are undergoing a radical shift, and usage statistics must take this into account; online journals are more than simply collections of articles, whereas the growth of open-access publishing means that journal articles are becoming more widely distributed with repositories and publishers hosting them.

To address these trends and ensure that usage statistics remain relevant, those of us involved in setting standards for the collection and recording of usage statistics, must, I think, address the following questions.

#### ***1. What Should We Measure?***

This is a Big Question, so let us cut it down to size by considering it in the context of journals. Online journals are much more than the electronic equivalent of print. In addition to articles, they offer supplementary data, video, and audio and also provide forums for interactive communication and the creation of online virtual communities. These new features are an

increasingly important part of the value offered by online journals, but they are not reflected in current usage statistics, which are still largely wedded to traditional measures of usage, such as “article downloads.” This must change, but before we prescribe new metrics, we need to have a deeper understanding of how these new features are used and decide what would be meaningful new measures. At COUNTER, we have stuck our toe into the water on this by providing usage reports that allow, for example, the usage of “non-textual items” such as video and audio, to be recorded and reported.

## ***2. How Do We Make it Easier for Librarians to Collect and Analyze Usage Statistics?***

Our experience with COUNTER has shown that librarians, on the whole, have a limited amount of time to devote to collecting and managing usage statistics. If the processes involved are too labor-intensive, usage statistics will not become an integral part of library management. This has been a problem, even for the COUNTER usage reports, which collect relatively simple usage data. The introduction of the SUSHI protocol (<http://www.niso.org/workrooms/sushi>), which allows the automation of the collection and consolidation of usage statistics, is an important step in making these processes more efficient, and further advances will be required on this front in the future.

## ***3. How Do We Make the Process Cost-Effective and Efficient for All Stakeholder Groups?***

We live in an age wherein our capacity to generate data is almost infinite, while our capacity to absorb and interpret data is limited by time and other priorities. Generating, collecting, processing, and storing data are costly processes, and all stakeholder groups involved must be convinced that the exercise is one in which the benefits justify the costs. The important thing here is to be highly selective in the usage data that we expect vendors to provide on a routine basis and to confirm that it does, in reality, provide librarians with new insights that justify the costs. As online publications mature, the features that are most valued will change and we must be prepared to modify accordingly the usage data that we collect.

## ***4. What Performance Metrics Can Be Derived from Usage Statistics?***

Many librarians already routinely calculate cost-per-download and cost-per-FTE, and these have become helpful barometers of the value, or

otherwise, of individual titles or collections of titles. There are other, usage-based metrics that are also being considered. Journal Usage Factor is one. ISI's journal Impact Factors, based on citation data, have become generally accepted as a valid measure of the quality of scholarly journals and are widely used by publishers, authors, funding agencies, and librarians as measures of journal quality. There are, nevertheless, misgivings about an over-reliance on Impact Factor alone in this respect. The availability of the majority of significant scholarly journals online, combined with the availability of increasingly credible COUNTER-compliant online usage statistics, raises the possibility of a parallel usage-based measure of journal performance becoming a viable additional metric. The U.K. Serials Group is sponsoring a research project to investigate the feasibility of developing such a measure (<http://www.uksg.org/projects>).

### ***5. How Do We Measure Usage in a Much More “Distributed” Online Publishing Environment?***

Librarians are no longer only “collectors” of content; the development of institutional repositories (IRs) means that they are also becoming distributors of content. IRs represent a considerable investment for the institutions concerned, and the management of these institutions will, I shall hazard, want to know something about the return they are receiving on that investment in terms of visibility, status, and user satisfaction; it is probable they will want some data on the usage of such facilities. There are now hundreds of institutional and subject repositories worldwide but no agreed standard for measuring usage. From COUNTER's perspective, it is clear that our current set of journal usage reports, which collect data at the journal level, do not meet the needs of repositories, which think in terms of individual items, such as articles, rather than in terms of collections of items, such as journals. Individual item usage reports are required, and a number of projects are already investigating how this can be achieved, including the PIRUS project, sponsored by JISC in the United Kingdom (<http://www.projectcounter.org/news.html>). It would be particularly desirable if publishers and repositories could adopt the same standards for the measurement of individual item usage. Only this will allow authors, their institutes, and the agencies that fund them to see a full, global picture of the usage of their publications.

These are, I think, the five broad questions we need to address as we look forward to the next steps for usage statistics for libraries. In doing so, we also need to acknowledge that there will always be a tension between the need for some stability in standards for usage statistics, so that vendors and libraries are not continually required to modify their systems, and the need

to keep up—or catch up, as some have put it!—with the latest technology. We must live in the real world.

It is time for another question and another request for answers from practitioners, academics, students, thinkers, and experts. It should not be a surprise to anyone that the next question is about the economy or, more specifically, how the current economic situation is impacting the field of what we call “electronic resource librarianship.” So, I ask you, “What impact will the recent economic downturn have on electronic resource acquisition, management, and/or use?” If you would like to share your answer to this question with *JERL* readers, send an email to [kingdp1@mailbox.sc.edu](mailto:kingdp1@mailbox.sc.edu), and your response will appear in a future column. Response length should be between 600 and 1,200 words. Also, if there is a question you would like to see posed in this column, please contact me.