THE
UNIVERSITY
OF RHODE ISLAND

University of Rhode Island
DigitalCommons@URI

Technical Services Department Faculty Publications

Technical Services

2008

E-Journal Usage Statistics in Collection Management Decisions: A Literature Review

Andrée J. Rathemacher
University of Rhode Island, andree@uri.edu

Follow this and additional works at: http://digitalcommons.uri.edu/lib_ts_pubs

Part of the Library and Information Science Commons

This is a pre-publication author manuscript of the final, published article.

Citation/Publisher Attribution

Rathemacher, Andrée J. "E-Journal Usage Statistics in Collection Management Decisions: A Literature Review." In *Library Data: Empowering Practice and Persuasion*, ed. Darby Orcutt, 71-89. Santa Barbara, Calif.: Libraries Unlimited, 2010.

E-journal usage statistics in collection management decisions: a literature review

By Andrée J. Rathemacher

Over the past several years, the online availability of scholarly journals has grown. Many research libraries now subscribe to journals in online format, either instead of or in addition to their print subscriptions. In the electronic environment, it is possible to measure how many times online journals are used. Statistics on usage can be requested from publishers or generated by local systems. These statistics can be used for collection management decisions, including whether to cancel or retain subscriptions. This chapter reviews the literature on employing usage as a factor in journal cancellation projects. It examines sources of e-journal usage data, issues in working with the data, the relevance of usage data of large e-journal packages, and the objectivity of the statistics themselves.

Usage-based journal cancellation and retention decisions [[[heading]]]

Journal use statistics have long been used by libraries in deciding which journals to cancel and which to retain. Medeiros reports that, "Goehlert, Maxin, and Broadus have described the value of usage statistics when making collection development decisions about print journals. Impetus for these studies of the late 1970s and early 1980s was the rising cost of serials, especially scientific/technological/medical (STM) journals" (2007, 234). Jaguszewski and Probst note that, "Usage and especially cost per use are central in determining whether a print title is a likely candidate for cancellation or withdrawal" (2000, 801). Kraemer adds, "For libraries with

expensive journal subscriptions, use data has been a mainstay for justifying the purchase of new journals and the cancellation of no longer cost-effective subscriptions" (2006, 164).

Usage as a proxy for more meaningful data [[[sub-heading]]]

Despite the widespread analysis of use data in making collection management decisions, many authors have noted that usage is a proxy for more meaningful but elusive data. According to Oliver Pesch, Chief Architect at EBSCO Publishing, "Statistics are a measurement of users' actions that we try to correlate to their intentions" (2004, 153). Peters adds, "What is usage, and what do we really know with confidence about how persons use information?...We need to be careful about any inferences we make from an analysis of usage data about the needs, interests, and preferences of users. How do users actually apply library resources and services to the information and learning projects of their lives?" (2002, 44).

Some authors have commented that the retrieval of an article (removing a journal issue from the shelf or successfully accessing a full-text article online) may not correspond to the article actually being read or cited at a later point in time. Best explains, "The library assumes that the article met the need, but for a heavy undergraduate population that may rely on the title only to judge an article's relevancy (often without looking at the abstract), students may not discover the inadequacy of the article until later in their research process" (2007, 205). Medeiros notes that "clearly there is some number of e-journal uses that merely meet a professorial requirement (e.g., a certain number of journal articles necessary in a bibliography) versus satisfying an information need" (2007, 240). According to Conger, usage measures an interaction with the material, but

beyond this, "the librarian has no way to interpret the reality of the user's expectation before the interaction, the degree to which that expectation was met after the interaction, or the outcome, or impact, of that interaction on the user's life as a researcher" (2007, 265)

There is also a question about whether all use is "created equal." Is the value of an article dependent on who is using it, or for what purpose? Boots and her coauthors state, "We do not know who within the organization is using any given journal, and whether it is used a little by many staff members or a lot by only a few staff. We do not know what value is attached to an article by the person who has downloaded it, and who is using it for their work" (2007, 197). Duy and Vaughan conducted a study at Concordia University Libraries in Montreal and found that "although reading an article and citing an article are different activities, and perhaps indicate different usefulness of an article, there is an overall correlation between journals that are looked at online, and those that are cited by local researchers." Nonetheless, they note that this might not be the case "at an academic institution where non-publishing students are presumably a large population of online journal readers" (2006, 516-17). Luther, in her heavily-cited 2001 white paper on electronic journal usage statistics, concludes, "Both publishers and librarians emphasize that measures of the level of activity do not indicate the value of an article. It is dangerous to assume that a popular title that is used by many students is worth more than a research title that is used by only a few faculty members working in a specific discipline. Other factors need to be considered" (2001, 123).

All in all, as Conger states, "Decisions that rely only upon usage data or inferential data suffer from a vacuum of experiential data about customer experience or customer outcomes" (2007,

265). Noonan and McBurney add, "Gathering usage data... will provide the staff only with quantitative data, which lends nothing to the assessment of a publication's quality or its intended use by the patron" (2007, 158).

Other criteria [[[sub-heading]]]

Perhaps because of the shortcomings of usage data, total use and cost per use are seldom the only factors considered in the cancellation or retention of periodical subscriptions. Enssle and Wilde examined the literature regarding what other information is used in decisions to cancel journals. While they found that "use is an integral part of the majority of serials cancellation projects," libraries also employ such measures as questionnaires and surveys, faculty journal preferences, interlibrary loan data, and citation and impact data (2002, 260). Other sources of information include subjective evaluation by subject specialists and checks against external benchmarks. Colorado State University Libraries also used Local Journal Use Reports (LJUR) from the Institute for Scientific Information (ISI), in which "ISI analyzes the information from its citation database and provides the institution with information about the journals in which their faculty have published or which their faculty have cited" (Enssle and Wilde 2002, 264).

At the Owen Science and Engineering Library (OSEL) at Washington State University, use and cost per use are the "primary evaluation factors for making retention decisions" (Galbraith 2002, 80). Yet "retention decisions are based on the following factors: use patterns over the last five years, population served, availability of other titles covering the same topic, impact factor, coverage in indexes owned by OSEL, availability in other local libraries, available document

delivery options, and faculty input" (Galbraith 2002, 85). Kraemer, at the Medical College of Wisconsin, "compiles an annual journal evaluation package that includes complete usage reports, low-use reports, high cost-per-use lists, a list of high-demand interlibrary loan (ILL) journals, and a list of requested journals." A Web-based survey of titles that were requested through interlibrary loan is then made available to all patrons (Bordeaux et al. 2005, 297). A survey was also used by the Lee Graff Medical & Scientific Library in Duarte, California; they designed a survey to combine with their usage statistics to determine what their users really wanted and needed (Wood 2006).

Enssle and Wilde conclude that, "On the surface, usage seems to be a good justification for canceling periodicals, however, upon closer examination, one soon discovers that this is not always the case...The primary problem with usage statistics is that they fail to take into consideration the impact of a journal in an individual subject area" (2002, 267). Best adds, "Usage data...is not and should not be the main driver for decisions on selection and retention. Selection and retention decisions...must be made with a proper understanding of the users being served, the users' needs, and the overall strength of the resources of the entire collection..." (2007, 200). According to Jaguszewski and Probst, "The librarian's knowledge of the collection, faculty research and teaching interests, user expectations, and current and future trends in research in related fields are equally important in ensuring that appropriate choices are made" (2000, 800).

While it is necessary to look at other factors and not make decisions based on usage statistics alone, usage has been and will remain an important criterion in cancellation and retention

decisions. As Gatten and Sanville acknowledge, "We now have a much more concrete measure of relative value, which is the number of uses (downloads). While sheer volume of use (i.e., cost-use analysis) is not the only measure of value, to fail to recognize use as the dominant starting point is to deny reality" (2004).

Sources of usage data for online journals [[[heading]]]

Vendor-supplied usage data [[[sub-heading]]]

Once a library decides to employ usage data as a factor on which to base e-journal cancellation and retention decisions, the library must decide whether or not to use vendor-provided or locally produced data. The majority of libraries rely on vendor-supplied usage data, which are readily available from a large number of publishers. In their 2004 survey of academic libraries in New Zealand, McDowell and Gorman found that 65% of libraries surveyed used vendor-supplied usage statistics to aid in collection management decisions. That number is sure to have increased significantly by 2008.

Previously, the literature was rife with complaints about the inconsistency of vendor-supplied usage data. McDowell and Gorman found in 2004 that the standardization of usage statistics across vendors through COUNTER or International Coalition of Library Consortia (ICOLC) guidelines had not yet been adopted by vendors. They found, for example, that there was no common definition of "use" between vendors: "Some vendors defined it as accessing the homepage of the e-resource, some as downloading the full text article, and others itemized use

counts of a number of different activities including searching or accessing table of contents, abstracts, or full text" (2004, 338).

Fortunately, since then the widespread adoption of COUNTER compliant statistics has largely solved these problems of inconsistency. COUNTER does a great deal to standardize these statistics and help make statistics between different publishers and platforms comparable. According to the COUNTER Web site, COUNTER was launched in 2002 in order to set "standards that facilitate the recording and reporting of online usage statistics in a consistent, credible, and compatible way."

A 2006 study by Duy and Vaughan concludes that COUNTER compliant statistics are a valid method to determine electronic journal use. Morrison adds that thanks to the efforts of initiatives such as COUNTER and the ICOLC Guidelines for Statistical Measures of Usage of Web-based Information Resources, "librarians are now beginning to see usage statistics based on these standards that are comparable across resources and platforms." Morrison acknowledges that a great amount of work still needs to be done, yet she concludes that "there are currently enough quality usage statistics available that this is now a factor in making financial decisions, such as the cancellation and retention of journals" (2007, 174).

Most libraries measuring the use of online journals will use the COUNTER Journal Report 1, "Number of Successful Full-Text Article Requests by Month and Journal." Boots and her colleagues at Cancer Research UK (CRUK) have "been persuaded that the number of full-text downloads is the most important statistic to measure" (2007, 189). They write, "CRUK has

developed a very simple model of cost-per-download by dividing usage statistics into journal costs. We have proposed to use this technique as part of the annual review of journal titles for cancellation or renewal" (2007, 187) Galbraith, at the OSEL at Washington State University, has concluded that "one use equals one access of a full text article" (2002, 81). In some cases, it is possible to examine other COUNTER reports, for example Database Report 3, which provides total searches and sessions on the vendor's platform, or the optional Journal Report 3, which offers additional measures including views of tables of contents, abstracts, and references (and which in COUNTER Release 3 will also measure requests of supplementary data sets, images, and videos). However, without a sophisticated statistics management system, so much data can lead to information overload and the inability to interpret any of it in a meaningful way.

If publisher does not provide statistics [[[sub-heading]]]

When relying on vendor-provided usage statistics for collection management decisions, one problem libraries encounter is the fact that some publishers do not provide usage statistics at all. Noonan and McBurney note that "there are a handful of small publications purchased individually that provide no statistics whatsoever" (2007, 156). Galbraith found that "most publishers who are not providing statistics are publishing just one or two e-journals" (2002, 85). Noonan and McBurney point out that the unavailability of usage statistics "makes it very difficult for the library to make informed collection development decisions" (2007, 156).

It may be to a publisher's advantage to not offer usage statistics when it comes to retaining subscribers; if a journal has low usage, it might get canceled, but if a journal does not report

usage at all, it is likely safe, at least from cancellations based on usage statistics. Peters speculates, "Some vendors may be reluctant to supply usage statistics, because they are fearful that low or uneven usage may cause libraries or consortia to demand changes in pricing, terms, and scope of content when license agreements come up for renewal" (2002, 43). Of course, it is conceivable that some libraries would consider the availability of COUNTER compliant usage statistics to be an important factor in deciding whether or not to subscribe to an online journal in the first place, or in whether to retain a subscription.

Locally generated usage statistics [[[sub-heading]]]

Instead of vendor-provided usage statistics, some libraries have chosen to rely on local statistics that measure links out to publisher sites through a journal A-Z list, an OpenURL link resolver, a proxy server, or in some cases, accesses to content hosted locally. Such practices have the advantage of providing statistics in cases where vendor-supplied statistics are not available, as discussed above. Kraemer explains, "For a shrinking number of electronic journals, locally gathered click-through counts from an A to Z list are the only statistics available" (2006, 166). For example, libraries that subscribe to Serials Solutions can access "click-through" statistics, which represent "any instance where a user clicks on a journal link within Serials Solutions. Thus, an institution can get a decent idea of how many times each database, each journal title within that database, or each journal title individually, has been accessed using the Serials Solutions interface" (Fowler 2007, 137). At the Medical College of Wisconsin, Kraemer has created a program that uses OpenURL linking data to track journal usage by year of publication, a measure that is not available in COUNTER compliant reports (Bordeaux et al. 2005, 296-7).

It is important to keep in mind, however, that not all researchers at an institution access subscribed resources through the library's interface, preferring to navigate directly to the journal or linking from emailed tables of contents or other sites (for example, Google Scholar). Therefore, these click-through statistics are not accurate. Ralston found a correlation between click-through statistics from a list of online journals and vendor-supplied statistics, yet he found that the "usage counts from the online journal list were not accurate enough to use with cancellation decisions" (2007, 51). He determined that "the correlation breaks down at the low use end of the online journal list statistics, precisely where it is needed to make accurate cancellation decisions" (2007, 60). Ralston further discovered that accesses from an A-Z title list could not be equated with articles downloaded, as his study "found an average ratio of 5.73 articles downloaded for each title access recorded on the list" (2007, 63). Kraemer offers a possible solution: "Since click-throughs represent only a fraction of the total use, that count is multiplied by a factor that has been derived from journals for which both click-through counts and publisher-provided COUNTER compliant data was available" (2006, 166).

There appears to be a consensus that vendor-provided usage statistics, for all their weaknesses, are the most complete and accurate data available. Fowler, in his study of Serials Solutions click-through data, found that "while the breadth and depth of the statistics it provides are impressive, they do not, as yet, serve as an alternative or replacement for the site statistics provided directly by journal providers, or journal platforms, such as Ingenta or MetaPress. They can, however, provide valuable adjunct statistics to complement and fill out these other statistics. In the cases where a particular journal title has no statistics provided by its platform or producer, Serials

Solutions may, in fact, be the only resource available for providing any usage statistics at all" (2007, 149).

Labor of gathering [[[sub-heading]]]

Regardless of their source, usage statistics for electronic resources take a great deal of time and effort to acquire and compile into a useful format. Despite their increased availability, vendor-supplied statistics have to be gathered, platform by platform. As Anderson explains, "The majority of vendors provide Web sites dedicated to usage statistics and, often, other administrative functions. This practice simplifies access to various data repositories, although the library must maintain a list of the access points for vendor statistics" (2007, 248). Weintraub rightfully complains, "Yale must maintain dozens of passwords to various databases in order to get usage statistics" (2004, 32). Fortunately, Release 3 of the COUNTER Code of Practice for Journals and Databases, effective January 2009, incorporates the SUSHI (Standardized Usage Statistics Harvesting Initiative) protocol. According to the COUNTER Web site, "Implementation of the SUSHI protocol by vendors will allow the automated retrieval of the COUNTER usage reports into local systems, making this process much less time consuming for the librarian or library consortium administrator." However, libraries will need an electronic resources management (ERM) system capable of fetching the statistics in order to take advantage of SUSHI.

Even after statistics have been gathered, they must be organized in such a way as to make them easily usable for decision-making purposes. Madeiros puts it well: "Despite the adoption of

COUNTER compliant usage statistics by a growing bloc of e-journal providers, the maintenance of usage statistics is a labor-intensive activity...Merely accumulating spreadsheets of e-journal usage, however, does not provide the kind of immediate and longitudinal analysis libraries need, in order to make both informed and on-demand collection management decisions. Still the practice of storing countless spreadsheets of usage data predominates among statistics maintainers, since no agreeable alternative presently exists" (2007, 236). Anderson adds, "Excel spreadsheets work well for cost-per-use analyses and for more sophisticated reports linking budget data and usage. Access and Oracle offer additional reporting capabilities, but...with a higher cost in time and personnel to develop that functionality" (2007, 251). Fortunately, a number of ERM systems now available offer an alternative to multiple spreadsheets for the storage and analysis of usage data..

Working with the numbers [[[heading]]]

Once usage statistics have been compiled, how should they be used and interpreted? How should the data be used to help make collection management decisions?

How many years of data to examine? [[[sub-heading]]]

One question to consider when making decisions based on usage data is how many years of data to examine when making cancellation or retention decisions. Some libraries, like the Owen Science and Engineering Library at Washington State University, calculate both an average and a

sum of all years of use available for each title (Galbraith 2002, 83). Other libraries compare data from the current year against past years. According to Peters, this answers the question, "How is usage of the e-resource evolving over time? After a year's worth of usage data have been compiled, the usage statistics for a month or quarter can be compared to those for the same month/quarter of the proceeding year" (2002, 44). Advanced comparisons or calculations will be easier with the use of an ERM or similar system.

In most cases, better decisions will be made by consulting multiple years of data. Comparing use over a few years will help avoid any aberrations in the data that occur during a limited time period. For example, a decrease in the use of a particular journal during a period of a few months to a year might be because a faculty member who made particularly heavy use of the title was on sabbatical, or has retired and not yet been replaced. On the other hand, a decline in use that is sustained over more than a year might reflect changing research interests in a department, in which case the journal should not be retained but perhaps traded for another title. Open lines of communication between librarians and researchers can help in the interpretation of these variations in usage.

Another factor to take into consideration is how long the library has subscribed to a particular online journal. In 1999, Townley and Murray wrote, "Usage rates are positively related to length of use. The longer a database has been available, the more likely it is to be used. A minimum of twelve to eighteen months seems to be necessary before heavy use will be observed. Libraries should commit to support any database for at least one year before assessing use" (1999, 38). Whether this is still the case is uncertain. As of 2008, it seems that journals and other electronic

resources with print counterparts are more likely to be used online than in print. For instance, Kraemer found in 2005 that "electronic usage has quadrupled in four years, while reshelving of print journals has dropped by more than half" (Bordeaux et al. 2005, 297). If there is no print alternative to an online journal, as is the case when libraries support only one format, online use might not take very long at all to reach a peak. Nonetheless, it is probably wise to take the amount of time a journal has been online into consideration when examining usage for cancellation purposes, giving more weight to usage from later years in cases where usage has grown over time.

Creating a benchmark [[[sub-heading]]]

Many libraries that rely on usage data for cancellation and retention decisions will rank their subscriptions by total use and/or cost per use and begin with a cut-off point, or benchmark, above or below which decisions will be made. In preparing to cancel journals, Enssle and Wilde at the Morgan Library at Colorado State University picked a "benchmark statistic to use for the initial cut. This benchmark statistic served to shorten the list of titles and eliminated the need to compile a complete set of statistics for each subscription in the collection." From there they compiled additional information for each title on the list of potential cancellations (2002, 266). Likewise, the Owen Science and Engineering Library (OSEL) at Washington State University maintains a database of in-house use statistics. Galbraith explains, "On a regular basis OSEL's selection librarians run a list of titles from the database that have received less than a designated number of uses during the year. The librarians then evaluate these titles in order to decide which to cancel and which to retain" (2002, 85).

Benchmarks used for cancellation decisions can be based on total use, cost per use, or both. In fact, these strategies were used long before usage statistics for online journals became readily available. In 1996, the Management and Economics Library at Purdue University developed a decision support system for a journals cancellation project based on duplication, price inflation, and use. Use was measured by scanning the bar codes of re-shelved journals for one semester, and titles were identified as candidates for cancellation if they were circulated or used in-house fewer than five times during the study (Nixon, 1996). Likewise, The Washington University School of Medicine in St. Louis looked at use of print, incorporating online use when available, and drew "a line... at a certain use level based upon the available funding" (Mercer 2000).

Benchmarks for cost per use are based on dollar amounts rather than times used. According to Hiott, "In 2003, the Houston Public Library used full-text units examined to measure the cost per use of its databases and to renegotiate prices for databases with high costs. With no information on what a unit should cost, the library set five dollars as an upper acceptable limit for a full-text article because the article cost for all but a few of its databases fell under this figure" (2004, 445). Galbraith at the OSEL uses a cost per use benchmark of $100 in addition to a benchmark for total use: "Use of a journal may be in the acceptable range, but the cost of the journal may make the cost-per-use unacceptable" (2002, 85). Sometimes the benchmark for cost per use can be based on an external factor. For example, Cancer Research UK compares cost per download "with the average cost of obtaining a copy of a journal article from document supply sources." (Boots et al. 2007, 183).

In the electronic environment, before canceling any title based on low use, the library should verify whether or not access to that title is working and that the content is accessible. If not, low use could be due to lack of access rather than lack of interest. Similarly, if a particular title or platform has had problems with connectivity during the year, this should be taken into account when interpreting the total usage for the time period.

Context needed [[[sub-heading]]]

One recurring theme in the discussion of benchmarks, however, is that they are inherently arbitrary unless placed into some kind of context. In 2001, before usage statistics were as prevalent, Luther commented, "Publishers are concerned that the data they share with librarians lack context. If, in the absence of such a context, usage data seem low, the publishers fear that librarians may use such information as a basis for canceling subscriptions" (2001, 121). Kraemer adds, "Complete as well as estimated or projected data requires a context before decisions can be derived from the data. Given the occasional contentiousness of journal budgeting negotiations, misrepresenting data unintentionally can have a lasting, damaging impact. The strength or weakness of data must be clearly identified" (2006, 164)

The need to place usage statistics into context by discipline is one of the key issues that has been raised. Luther writes, "Known differences in information-seeking behavior among users in various scientific disciplines warrant additional study to identify usage patterns. As more data are examined on use and behavior, it may be possible to establish average levels of use for different subject areas or user groups" (2001, 123). Enssle and Wilde agree that "when examining the

statistics, it quickly becomes evident that there is no uniform standard for 'low-use;' it is best to compare journals within individual disciplines rather than across the board. In some cases 'low-use' journals had been used fewer than twice a year, while for other disciplines 'low-use' journals had been used fewer than 15 times a year. Intra-disciplinary comparisons help ensure that the cancellations are more evenly spread among the departments, and that the collection will not lose its diversity" (2002, 267).

Another context suggested to help make sense of usage statistics is that of peer institutions. Peters writes, "Just as there are basic ways to analyze e-resource usage statistics, so too are there basic contexts into which these usage reports can be placed...Usages statistics for the same e-resource from peer institutions can be compared" (2002, 44). Best concurs: "there is no universal consensus as to what constitutes acceptable or adequate amounts of usage. Without access to data for peer institutions with similar collections, it is doubtful that we will ever be able to provide proper benchmarks for usage rates" (2007, 200). At least one vendor, Serials Solutions, already provides benchmarks of click-through usage data by peer institutions. Fowler explains that subscribers can use benchmarking data available from Serials Solutions to compare and contrast the median statistics at their institutions with those at peer institutions that are also Serials Solutions subscribers. Peers are divided into categories by library type (academic, public, and special). Public libraries are then broken down by number of residents served, and academic libraries are grouped by Carnegie class (2007, 143-46).

Other options for creating context have been proposed as well. Luther suggests the possibility of benchmarking by journal price or number of articles available. She writes, "To base comparisons

on the use of large or very popular journals (e.g. *Nature*, *Science*, or *Cell*) sets an artificially high benchmark for other titles with fewer articles available for use. This raises the question of whether the measure of activity should be relative to another factor, such as the price of the journal or the number of available articles, which puts the measure in a context" (2001, 123). Finally, some have suggested that e-resource usage statistics would be more meaningful if placed in the context of institutional population. Relating use levels to the number of potential users would allow for an examination of the growth in use over time of a library's e-resources by its users (Blecic et al. 2001, Peters 2002.

Titles with very low or no use are not unusual [[[sub-heading]]]

In relation to benchmarks, it should not come as a surprise when a library finds that a substantial number of their subscriptions receive very low or even no use. In 1969, Trueswell posited that a well known rule of inventories, the "80/20 Rule," applied to libraries, in that 80 percent of circulation requirements are satisfied by 20 percent of a library's holdings. Trueswell's findings have been reproduced in subsequent studies. In their 2000 study of the usage of 194 titles available both in print and online at the Norris Medical Library at the University of Southern California, Morse and Clintworth note that "the concentration of usage on the most popular titles was almost identical for the print and electronic lists. In both cases just 20% of the titles accounted for nearly 60% of the usage, and, conversely, the bottom 40% of both ranked lists accounted for only 9% of total usage." Kraemer at the Medical College of Wisconsin finds that "the highest use journals make up eighty percent of the total journal use" (Bordeaux et al. 2005, 297). Galbraith at the Owen Science and Engineering Library writes: "OSEL data collection has

shown that it cannot be assumed that all electronic journals are heavily used. The database has revealed some titles that have received no uses in the time that they have been available" (2002, 88). Morse and Clintworth (2000) sum up the issue when they conclude that studies like theirs indicate "that the large spread in usage levels between titles, which librarians have long observed in the print domain, is being duplicated in the electronic one."

Reflections on publishers and packages [[[heading]]]

One situation that results in a lack of meaningful usage data, at least for cancellation purposes, is the bundling by publishers of a large number of their journal titles into packages, often referred to in the literature as "portfolio packages" or the "big deal." While the price of the journals included, especially of those journals a library already subscribes to, is often used to calculate the initial cost of a portfolio package, most packages do not subsequently break down the package cost by title. Even if they do, the ability of a library to cancel individual titles in a package is usually contractually curtailed.

Because these packages are priced at a discount, costing much less than the combined list price of all included titles, it is usually the case that the aggregate cost per use for the entire package is low compared to non-packaged e-journals. The lower cost per use suggests that these packages are a good value overall. However, critics of these packages, most notably Frazier (2001, 2005), have written in depth on the hazards of the "big deal," which can be summarized as follows:

- Libraries lose much of their ability to shape their collections according to local needs. "Big deal" packages monopolize library budgets, preventing money from being used for

other resources, for example journals from smaller publishers and the "big deal" publisher's competitors.

- Libraries have less flexibility in response to budget cuts. In order to avoid the disruption that would be caused by canceling the portfolio package, they will instead cancel journals from smaller, not-for-profit publishers. This punishes the small publishers, which are for the most part not responsible for the crisis in serials costs, while favoring the large, for-profit publishing corporations that raise prices indiscriminately. As stated by Jaguszewski and Probst: "The prospect of canceling journals published by learned societies and professional associations in order to support lesser-used titles from commercial publishers is one most librarians will want to avoid" (2000, 804).

- Diversity among library collections decreases, as many libraries subscribe to the same packages.

- These trends in turn enhance the market power and market share of the large "big deal" publishers.

Regardless of one's stance toward these packages, it is hard to argue with Best's contention that "aggregators do bundle their weaker titles with their strongest titles. As a result, libraries end up paying for titles they neither want nor need" (2007, 206). It is likely for these reasons that the Committee on Institutional Cooperation (CIC), a consortium of 12 research universities (the 11 members of the Big Ten Conference and the University of Chicago), recommended in 1998 that, "Bundling of titles as a provider's 'all-or-nothing' electronic journal product is not the best solution for all libraries and should not be the sole model for licensing these titles. Title-by-title

purchase may be preferred and this option should be available," and "Cancellation of electronic journals should be possible on a title-by-title basis."

While many "big deal" packages offer COUNTER compliant usage data, these data are of limited use to subscribers. Although usage statistics are provided on a title-by-title basis, restrictions on canceling individual titles mean that such data serve only to alert the library to journals that are either more or less used. If a large number of journals in the package do not receive much use, the library may consider canceling the package or attempting to renegotiate the price at the end of the contract period.

Cost per use data for these packages is also less useful than similar data for individually subscribed e-journals. Meaningful cost per use data can only be calculated for the package as a whole, since the relative cost of the journals in the package is not known, although some libraries (and ERM statistical software) divide the total package cost by the number of titles to determine an average cost per title. As Boots and her coauthors explain, "the cost of individual journals cannot easily be separated from the overall total. In these cases, the total cost is divided equally across the number of titles in the deal" (2007, 191). In any case, as Noonan and McBurney state, "Usage statistics availability does not mean the library can always utilize them if it has purchased bundled packages, or if the library is locked into a multiyear deal" (2007, 156).

Not as objective as they appear [[[heading]]]

Although usage statistics are not the only criterion on which to base cancellation decisions, at first glance they appear to have the advantage of being straightforward, objective measures that reflect the value that a given institution's users place on a journal. However, a closer look reveals that usage statistics are not as simple as they seem; in fact they could easily be characterized as "dirty data." There are many factors that can distort usage statistics or make comparisons between them invalid. While most of the problems identified below relate specifically to vendor-supplied usage statistics of online journals, some are equally relevant when using locally-generated online journal statistics or in the print environment.

Title changes [[[sub-heading]]]

One problem in working with usage statistics is that publishers are inconsistent about how they handle title changes, split titles, and merged titles. Are the usage statistics for a particular journal consolidated under the new title of a journal, divided between new and old titles based on the date of the article downloaded, or duplicated for each title? To accurately compare journals by use, usage data for current and former titles of a journal must be merged to reflect total use, otherwise there is a risk that use for the new title will be artificially low. With a large subset of journals, this is a daunting task. It may not be necessary to consolidate the usage data for all changed titles on a regular basis, although some libraries do (see Galbraith 2002 and Wulff and Nixon 2004), and ERM systems can help in this task. Certainly, however, before a journal is canceled based on low use, it should be checked for a recent title change on the chance that use has accrued on the basis of the previous title. Henle and Cochenour sum up the problem with statistics when journals change titles, split into multiple parts, merge, or change publishers: "Any

of these issues produces fractures in a journal's usage statistics that may go unnoticed and, thus, produce a false picture of the actual importance of a title for the local collection. While such detailed statistics are valuable, compiling them to ascertain the use of a given subscription and/or subject becomes a complicated affair. The task of integration across titles...is too manually cumbersome to justify routine processing and is better saved for specific requests" (2007, 18-19).

Bundled titles [[[sub-heading]]]

Some e-journal subscriptions are bundled such that a subscription to one title includes a small number of additional titles. This is on a much smaller scale and not at all comparable to the portfolio or "big deal" packages discussed above. For example a 2008 subscription to Wiley-Blackwell's *International Studies Quarterly* includes four other titles from the International Studies Association: *International Studies Perspectives*, *International Studies Review*, *International Political Sociology*, and *Foreign Policy Analysis*. These additional titles are not priced individually or even available as separate subscriptions; they are simply included with a subscription to the "mother" title.

As is the case with "big deal" packages, the total number of successful full-text article requests will be available on a title-by-title basis, but calculating cost per use for each title is complicated. Libraries have to decide whether to divide the cost of the bundle equally between all the titles and then calculate cost per use for each title based on this average cost, or whether to perform a separate calculation for the package as a whole, with an aggregated number of article requests

divided into the total price. Librarians should also think about whether or not it is important to apply the same methodology to all such bundles. Is it important to be consistent? Or should this decision be made on a case-by-case basis? For example, if all of the journals in a bundle are desired subscriptions, the value of the bundle could be calculated as one unit. In another case, perhaps the library has determined that only one title in a package is important to the collection. In that instance, cost per use could be calculated for the important title only, while the data for the other titles is ignored.

Either way, these small bundles can easily be overlooked, resulting in inaccurate cost per use calculations. Since the price for the bundle is actually the price of the mother title, it is easy to accidentally ignore the child titles and calculate cost per use of the mother title based on the use of the mother title alone. This makes each use of the subscription appear more expensive than if use for all titles were included in the calculation. If cancellations are being made based on cost per use, the mother title (along with the children) might be slated for cancellation based on this inflated cost per use number.

Multiple platforms [[[sub-heading]]]

Electronic access to a given journal is often available to an institution's users through multiple platforms, and this content may or may not reflect more than one subscription. For example, a subscription to the current volume of a journal might be available on the publisher's own platform as well as on IngentaConnect, and EBSCOhost EJS. To accurately sum up total uses, or to calculate cost per use based on the current subscription cost, requires that the total uses on

each of these platforms be added together. To further complicate matters, the same title may have back volumes available through JSTOR and/or a full-text indexing and abstracting database, for example EBSCOhost's Academic Search Premier. For the purposes of measuring total use of the title, the use on these platforms must be included. However since JSTOR and Academic Search Premier are separate subscriptions, the usage from these platforms should not be included in calculating cost per use of the current volume using the annual subscription fee of the journal. Multiply this situation by the number of titles a library accesses online and the scope of the problem becomes clear. Furthermore, usage from JSTOR and full-text databases may actually cannibalize some of the use from the publisher's platform, making usage from the publisher's platform appear lower and inflating cost per use.

Many librarians have struggled with these problems. Luther writes, "Constructing a complete picture of use is further complicated by the existence of journals in multiple formats that are available through multiple sources or distribution channels" (2001, 123). Boots and her colleagues faced this problem when Cell Press switched from their own platform to Science Direct and the usage for the same time period was reported differently for each platform. It was not clear whether to combine the statistics or if that would be double counting. They note, "Problems occur when a publisher launches a new Web site, or moves to a new aggregator platform. This frequently requires a complete change in the method whereby usage statistics are counted, and the format in which they are displayed. If this occurs midyear, it becomes very difficult to merge the new usage data with the old" (2007, 190).

Amount of content available online [[[sub-heading]]]

While missing or double counted usage statistics resulting from title changes, bundling, or platform variations can impact the validity of usage statistics, with enough effort the data can be adjusted to account for these anomalies. A more serious shortcoming of vendor-supplied online usage statistics as they currently stand is the varying number of years of content available online for different titles, which makes accurate comparisons between titles based on use impossible.

To illustrate: if articles from Journal A are available online back to 1950, while articles from Journal B are available online only for the last ten years, it is not surprising that total usage for Journal A will be higher than for Journal B, with related effects on cost per use, even though researchers may be using articles from current issues of both journals at the same rate. It is interesting to speculate whether the increased use that results from access to more online content motivates in part the common practice of publishers providing online access to the latest ten years or so of a journal, even if a library has subscribed in online format for a much shorter time. This might also be one motivating factor behind the move by many publishers to digitize their backfiles.

Luther writes, "A collection becomes more useful when the amount of archival content available online increases, especially if it is well indexed. When backfiles are included with the current subscription or basic service, the user has more articles to view, and this will affect usage" (2001, 124-25). Medeiros points out that, since cost per use is calculated with the cost of one subscription year only, the "calculation is inherently flawed..., since publisher-provided statistics do not reveal the percentage of ... uses that were of [the subscription year] journal issue content

only" (2007, 237). Holmström (2004) makes a similar point: "Currently ROI [return on investment] of 2003 expenditures is calculated relating the money invested to gain access to the journals published in 2003 to the total number of downloads in 2003 from the publisher's collection for all years. This way ROI appears to be larger than it actually is by including downloads that should not be included. However, the current method also neglects future downloads to older articles and therefore makes ROI appear smaller than it actually is." Usage studies of print journal collections often suffer from the same flaw (Galbraith 2002, 81), yet in many cases it is possible to track use by publication date for print journals by scanning barcodes of bound volumes or counting use of unbound issues only.

A number of solutions to this problem have been suggested. Many have called for publishers to provide usage statistics not only by date of use, but by date of publication. McDowell and Gorman found in their survey of New Zealand academic libraries that the least commonly provided vendor usage statistic in responders' libraries was "publication date of material accessed," yet librarians identified this as important (2004, 336). Pesch adds, "When a library can determine how frequently various subsets of their user community are accessing full-text based on age of articles, they can make more informed decisions on the relevance of subscribing to backfile collections versus paying for on-demand access" (2004, 150).

Providing monthly and annual usage data for each journal by year of publication would add a great deal of complexity to the data, yet it would allow for the number of uses of the current volume to be separated out. This would enable a more accurate calculation of cost per use and

basis for comparing total use between journals, especially in relation to decisions about canceling or retaining current subscriptions.

A step in this direction will come with COUNTER Release 3, to be implemented in 2009. COUNTER Release 3 is slated to require a new report, "Journal Report 1a: Number of Successful Full-text Article Requests from an Archive By Month and Journal" (currently an optional report in Release 2). Journal Report 1, "Number of Successful Full-Text Article Requests by Month and Journal," will still include all use of a journal. By subtracting use of an archive from all use, an estimate of current, or "non-archive" use could be made. While not precise, this would at least remove the effects of backfiles from the usage data of some journals. One problem is that the new report is required only of vendors "who provide separately purchasable journal archives" (COUNTER). There is also no uniform definition of exactly what an archive is. More helpful would be a report that separated out the use of material published a given number of years ago (for example, five years).

Indeed, although a breakdown of usage statistics by a fixed time period is not currently available, it would go a long way toward making usage statistics more comparable. It would then be possible to calculate and compare the use of the latest five years of content evenly across all journals. Five years is a good cut-off point, since a study of monthly publications by Maxfield and her colleagues found that "use peaks at three months after publication. By eighteen months after publication, fifty percent of use has occurred, and by five years after publication, ninety percent of use has occurred" (1995, 76). Holmström (2004) cites similar data from Tenopir and King that breaks down percentage of articles read by how recently they were published and

suggests that this can be used to estimate the number of uses by year of publication. For example, according to Holmström's adaptation of data from Tenopir and King, 58.5 percent of all readings of scholarly articles by university scientists occur for articles that are one year old or newer. By multiplying total use by 58.5 percent, one could estimate the amount of that use which occurred in the current year, and thus match current use with current subscription price to better calculate cost per use. Medeiros makes a similar claim: "Given the limitation of current statistics-reporting capabilities, cost-per-use should be derived from a Tenopir-King-like formula, which factors in all uses of an e-journal over the lifetime of that e-journal" (2007, 238). Yet such an approach assumes the presence of available backfiles for all online journals, which is not the case. Holmström (2004) acknowledges this when he states, "the reliability of these estimates is affected by the lack of backfiles… In order to accurately measure ROI, we need download statistics by time period published" (Holmström 2004).

Others have suggested a slightly different approach: to create a ratio comparing usage of a journal in a given time period to the total number of articles available online from that journal. Luther wonders "whether the measure of activity should be relative to another factor, such as the price of the journal or the number of available articles, which puts the measure in a context" (2001, 123). Hahn and Faulkner elaborate on this idea by introducing a metric called "content adjusted usage," which is calculated by dividing the number of full-text accesses by the total number of articles online for that journal. This is a way to "compare the usage of journals that offer widely differing numbers of articles online" (2002, 218-19). To calculate "content adjusted usage" would require publishers to provide data on total number of articles available or lengthy studies to count or estimate total articles, and such data would need to be updated on a regular

basis. Whether to make the calculation using total articles available online or only the number of articles available to an institution is another issue to consider. All in all, without more data from vendors as well as the assistance of staff with programming skills to create a database to store and manipulate this kind of data, a "content adjusted usage" calculation, while useful, would be extremely difficult and time-consuming to produce on a large scale.

In the face of this problem, some libraries have given up on vendor statistics altogether and use data from their local systems. For example, Kraemer, at the Medical College of Wisconsin, has resorted to using OpenURL data to track journal usage by year of publication (Bordeaux et al. 2005, 296-97). The problem with this approach, besides requiring technical expertise in tracking and converting OpenURL data into a tally of uses by journal and year, is that not all researchers access the library's resources using the library's link resolver; many researchers navigate directly to the journal or use services like Google to identify articles. Kraemer, too, believes that requiring COUNTER reports to group usage by year of publication would be the best solution (Bordeaux et al. 2005, 297).

Technical issues [[[sub-heading]]]

Even if usage statistics as provided by vendors were normalized for title and platform changes and were broken down by year of publication, there are other behind-the-scenes ways in which usage statistics may fail to record the true use of a journal.

Full text delivered during article-level linking [[[sub-sub-heading]]]

The measure of successful full-text downloads may be inflated if HTML full-text is displayed with the initial citation display or when a user is directed to an article through a linking mechanism (e.g., CrossRef). Oliver Pesch, Chief Architect at EBSCO Publishing, explains, "If a user is merely browsing the 'detailed' display of a list of articles, they may well be accumulating the full text retrieval count by virtue of this being shown automatically. The user may not even have cared about the article, but the count is still accumulated" (2006, 153). After viewing the HTML, the user may then view, print, and/or e-mail the PDF version, and these actions may be counted as additional uses (Pesch 2004, 148). A study by Davis and Price in 2006 found the ratio of PDF to HTML views to be inconsistent across publisher interfaces, even after controlling for differences in publisher content, to be inconsistent across publisher interfaces, which may have been the result of this phenomenon.

COUNTER Journal Report 1, "Number of Successful Full Text Article Requests by Month and Journal," provides separate year-to-date totals for HTML and PDF articles requested. This allows for any large discrepancies between HTML and PDF to be noted. As Pesch explains, "By separating out HTML and PDF, the institution can make better informed decisions related to a given journal" (2006, 153). However, there is still some confusion on how best to handle the data. JSTOR, for example, only has articles in PDF format, and counting only HTML would exclude them. As Hiott asks, "Does a library add together PDF and HTML articles examined, and how much double counting does that represent?" (2004, 450). It is probably best to use total articles downloaded without making a distinction between PDF and HTML, with the understanding that the usage of some journals may be inflated if the full text of any format is set to display automatically when linking to the article. Perhaps as a result of the awareness of this

problem, many of the major publishers and platforms have recently set their article-level links to display only a citation, or a citation with abstract and/or list of works cited, with links to full text in HTML or PDF format. While not providing one-click access to full-text might inconvenience the user, it certainly results in more realistic usage statistics.

Federated search engines, robots, and crawlers [[[sub-sub-heading]]]

Metasearch or federated search products can dramatically affect usage statistics, primarily for searches and sessions as opposed to article requests. However, full-text articles requested can also be inflated if the full text is retrieved automatically with the search results. Hits by internet robots and crawlers, as well as LOCKSS systems, can also be registered as full-text article requests. Fortunately, the forthcoming COUNTER Release 3 includes "new protocols that require activity generated by internet robots and crawlers, as well as by LOCKSS and similar caches, to be excluded from the COUNTER reports" (COUNTER). Vendors must comply from the beginning of 2009.

Pre-fetching and double-clicking [[[sub-sub-heading]]]

Other technical issues that have been addressed by COUNTER are pre-fetching and double-clicking. Pre-fetching occurs when Google or other services improve performance by automatically fetching the full pages for the first few results of a search. As Pesch explains, "When the user does click the link, another header request is made, that could be considered yet another full text request. Without some kind of control, this activity could result in significant

over-counting" (2006,159). COUNTER Release 3 should fix this issue, as it contains "a new protocol that requires prefetched full-text articles to be excluded from the figures reported in all reports containing full-text requests" (COUNTER).

Double-clicking by impatient users with slow connections can also increase the tally of full-text requests. Different publishers may filter multiple clicks according to different rules, which can affect usage. According to the Joint Information Systems Committee (JISC), the COUNTER Code of Practice has specified a filter to eliminate this problem. In sum, COUNTER compliant statistics should no longer be affected by pre-fetching or double-clicking, however librarians should be aware of these issues when working with non-COUNTER compliant usage statistics.

Effects of site interface [[[sub-sub-heading]]]

Many authors have commented on the potential of the design and functionality of the vendor's Web site to influence usage. Davis and Price studied general levels of usage across journal publisher platforms and compared the use levels of identical content on different platforms, finding that "the design of a publisher's electronic interface can have a measurable effect on electronic journal usage statistics" (2006, 1243). Kraemer maintains that "electronic journal use is more dependent on the range of access mechanisms available, than was the typical use scenario for print journals. If, for example, the access infrastructure for electronic journals is poorly developed, the data will show use levels well below their potential" (2006, 171).

Most of the effect of interface on usage levels appears to center around convenience. Usage will be higher as ease of access increases, while additional steps in retrieving content will cause use to be lower. In explaining the popularity of electronic access over print in the late 1990s, Morse and Clintworth (2000) concluded that "there is no question that a substantial percentage of the electronic journal usage reported here is attributable to the convenient full-text links provided in the heavily-used Ovid databases. User preference for electronic versions of articles that are not similarly linked to popular databases would be much less pronounced than that observed in this study." Luther comments, "the user's experience of the interface also will significantly affect the results. Both Academic Press and the American Institute of Physics (AIP) noted that they experienced surges in usage after they introduced new platforms that simplified navigation and access" (2001, 125). Wulff and Nixon find that researchers show "the same relative behavior in the electronic environment as in the print environment, especially when multiple steps [are] not required to locate the electronic journal. One-click electronic access from citation to article is important to users and an important factor in electronic journal usage. At times, users find that the best article is the most convenient article, and they choose to end their search at that point" (2004, 321).

Publisher manipulation [[[sub-sub-heading]]]

Value-added features on a site can also contribute to higher use that might otherwise be the case. In a review of journal use data at the Medical College of Wisconsin Libraries, Kraemer found that "the rate of increase in use for individual electronic journals appears to show higher rates for those journals with advanced linking features, e.g. linked references, OpenURL support, and

other features that 'enhance the presence' of a journal in an electronic environment" (2006, 169-70). It is interesting to note that this phenomenon, if true, could privilege the content from large, commercial publishers with substantial resources to devote to interface design over smaller publishers with less sophisticated platforms. Davis and Price acknowledge that "it is entirely possible for a publisher to optimize its interface to maximize the total number of article downloads. In an environment that requires justifying price per download, this may be an understandable goal" (2006, 1247).

Numerous authors have raised the related concern of whether publisher-provided usage statistics are inflated. Luther writes, "Publishers who make usage data available are aware that this information will be used to assess the value of their journals. Consequently, they want to ensure that usage is high so that the cost per-use is low compared with that of other publications. Publishers and librarians with experience in electronic databases agree that marketing to users - whether librarians or library patrons - and making them aware of the availability of the resource and its features have a noticeable impact on usage" (2001, 124). A common way in which publishers are able to market to researchers is through encouraging them to create accounts on the publisher's site by registering their e-mails. This allows users to opt in to added features (for example, the ability to save searches or sign up for e-mail tables of contents for their favorite journals), which are likely to increase use.

In a study at North Carolina State University, Duy & Vaughan (2003) found evidence that publisher-provided usage statistics overstated use in comparison to in-house data. The researchers compared journal use as measured by library Web server logs to vendor-supplied

data for a twelve-month period. Fewer than half of the library's vendors were reasonably close to the in-house use data. However, it is not safe to assume that all users are connecting through the library's site. Many users may navigate directly to the vendor site to retrieve content; thus, it makes sense that vendor-supplied data would show higher use. Nonetheless, as Medeiros states, "It is not hard to imagine some publishers padding actual use as a way of enticing renewal of certain middle-tier journals. Perhaps in response to this concern, COUNTER employed a rigorous auditing requirement to ensure the authenticity of statistics" (2007, 239).

When canceling journals based on usage data, therefore, using COUNTER compliant data is essential to ensure a level playing ground across publishers. While value-added Web site features may encourage researchers to make greater use of a publisher's content, with COUNTER compliance we may assume that use is *bona fide*. Since cancellation decisions are made based on *low* usage (either as a direct measure or by calculating cost-per-use), publisher attempts to inflate usage would be more likely to result in an under-utilized journal being retained than in a heavily-used title being canceled based on low use.

Conclusion [[[heading]]]
This chapter has reviewed the literature related to the use of electronic journal usage statistics in cancellation and retention decisions. It found that usage is largely a proxy for more meaningful information that is difficult to measure, and that most libraries base collection management decisions on other factors in addition to usage. It explored the sources of usage data for electronic journals (retrieving statistics from a vendor and generating them inhouse) and showed that each method is labor intensive and notes that ERM systems under development will assist

with these tasks. Questions related to working with the data were addressed: how many years of usage data to examine, the need to create benchmarks for decision-making, and the desirability of placing the data in a larger context (for example by discipline or peer institution). Also examined was the affect of the bundling of e-journals into large packages on the ability to make decisions based on usage levels. Selected characteristics of usage statistics were revealed that might require adjustment of the data or at least care in its interpretation. These include the effects of title changes, the bundling of titles, the availability of journal content on multiple platforms, the number of years of content online, and technical issues that might result in inflation in usage. Throughout, the importance of the COUNTER standard in ensuring reliable and comparable data was emphasized. Many of the difficulties of working with usage data, as well as shortcomings in the data itself, will likely be resolved by evolving standards, including COUNTER, and by the evolution of ERM systems for tracking and manipulating data.

References [[[heading]]]

Anderson, Elise. 2007. The next steps in developing usage statistics for e-serials. In *Usage statistics of e-serials*, ed. David C. Fowler, 245-60. Binghamton, NY: Haworth Information Press.

Best, Rickey. 2007. Lies, damn lies, and usage statistics: What's a librarian to do? In *Usage statistics of e-serials*, ed. David C. Fowler, 199-214. Binghamton, NY: Haworth Information Press.

Blecic, D. D., J.B. Fiscella, and S.E. Wiberley, Jr. 2001. The measurement of use of web-based information resources: An early look at vendor-supplied data. *College & Research Libraries* 62, no. 5: 434-53.

Boots, Angela, Julia Chester, Emma Shaw, and Chris Wilson. 2007. E-journal usage statistics in action: A case study from Cancer Research UK. In *Usage statistics of e-serials*, ed. David C. Fowler, 183-198. Binghamton, NY: Haworth Information Press.

Bordeaux, Abigail, Alfred B. Kraemer, and Paula Sullenger. 2005. Making the most of your usage statistics. *Serials Librarian* 48, no. 3/4: 295-99.

CIC Collection Development Officers and CIC Electronic Resources Officers. 1998. Preferred practices for CIC licensing of electronic journals. Committee on Institutional Cooperation (CIC) Center for Library Initiatives.
http://www.cic.uiuc.edu/programs/CLIConsortialAgreementProgram/archive/BestPractice/Prefer edPractices.pdf.

Conger, Joan E. 2007. Usage statistics in context: Develop effective assessment practices through collaboration. In *Usage statistics of e-serials*, ed. David C. Fowler, 261-73. Binghamton, NY: Haworth Information Press, 2007.

COUNTER: Counting Online Usage of Networked Electronic Resources. http://www.projectcounter.org/.

Davis, Philip M., and Jason S. Price. 2006. eJournal interface can influence usage statistics: Implications for libraries, publishers, and Project COUNTER. *Journal of the American Society for Information Science & Technology* 57, no. 9: 1243-48.

Duy, Joanna, and Liwen Vaughan. 2006. Can electronic journal usage data replace citation data as a measure of journal use? An empirical examination. *Journal of Academic Librarianship* 32, no. 5: 512-17.

Duy, Joanna, and Liwen Vaughan. 2003. Usage data for electronic resources: A comparison between locally collected and vendor-provided statistics. *Journal of Academic Librarianship* 29, no. 1: 16-22.

Enssle, Halcyon R., and Michelle L. Wilde. 2002. So you have to cancel journals? Statistics that help. *Library Collections, Acquisitions, & Technical Services* 26, no. 3: 259-281.

Fowler, David C. 2007. Serials Solutions and the art of statistics gathering. In *Usage statistics of e-serials*, ed. David C. Fowler, 129-49. Binghamton, NY: Haworth Information Press.

Frazier, Kenneth. 2001. The librarians' dilemma: Contemplating the costs of the "big deal." *D-Lib Magazine* 7, no. 3 (March), http://www.dlib.org/dlib/march01/frazier/03frazier.html.

Frazier, Kenneth. 2005. What's the big deal? *Serials Librarian* 48, no. 1/2: 49-59.

Galbraith, Betty. 2002. Journal retention decisions incorporating use-statistics as a measure of value. *Collection Management* 27, no. 1: 79-90.

Gatten, Jeffrey N., and Tom Sanville. 2004. An orderly retreat from the big deal: Is it possible for consortia? *D-Lib Magazine* 10, no. 10, http://www.dlib.org/dlib/october04/gatten/10gatten.html (accessed April 20, 2008).

Hahn, Karla L., and Lila A. Faulkner. 2002. Evaluative usage-based metrics for the selection of e-journals. *College & Research Libraries* 63, no. 3: 215-27.

Henle, Alea, and Donnice Cochenour. 2007. Practical considerations in the standardization and dissemination of usage statistics. In *Usage statistics of e-serials*, ed. David. C. Fowler, 5-23. Binghamton, NY: Haworth Information Press.

Hiott, Judith. 2004. Collecting and Using Networked Statistics: Current Status, Future Goals. *Library Quarterly* 74, no. 4: 441-454.

Holmström, Jonas. 2004. The return on investment of electronic journals: It is a matter of time. *D-Lib Magazine* 10, no. 4, http://www.dlib.org/dlib/april04/holmstrom/04holmstrom.html (accessed April 20, 2008).

Jaguszewski, Janice M., and Laura K. Probst. 2000. The impact of electronic resources on serial cancellations and remote storage decisions in academic research libraries. *Library Trends* 48, no. 4: 799-820.

Joint Information Systems Committee (JISC). COUNTER Filter: Improving the Comparability of Usage Statistics. Joint Information Systems Committee. http://www.jisc.ac.uk/whatwedo/programmes/programme_pals2/project_counter.aspx (accessed May 14, 2008).

Kraemer, Alfred. 2006. Ensuring consistent usage statistics, part 2: Working with use data for electronic journals. *Serials Librarian* 50, no. 1/2: 163-72.

Luther, Judy. 2001. White paper on electronic journal usage statistics. *Serials Librarian* 41, no. 2: 119-48.

Maxfield, Margaret W., Rebecca DiCarlo, and Michael A. DiCarlo. 1995. Decreasing use of monthly serials after publication date. *Serials Librarian* 27, no. 4: 71-77.

McDowell, Nicola, and G. E. Gorman. 2004. The relevance of vendors' usage statistics in academic library e-resource management: A New Zealand study. *Australian Academic & Research Libraries* 35, no. 4: 322-43.

Medeiros, Norm. 2007. Uses of necessity or uses of convenience? What usage statistics reveal and conceal about electronic serials. In *Usage statistics of e-serials*, ed. David C. Fowler, 233-43. Binghamton, NY: Haworth Information Press.

Mercer, Linda S. 2000. Measuring the use and value of electronic journals and books. *Issues in Science and Technology Librarianship* 25, http://www.istl.org/00-winter/article1.html (accessed March 27, 2008).

Morrison, Heather, 2007. The implications of usage statistics as an economic factor in scholarly communications. In *Usage statistics of e-serials*, ed. David C. Fowler, 173-82. Binghamton, NY: Haworth Information Press.

Morse, David H., and William A. Clintworth. 2000. Comparing patterns of print and electronic journal use in an academic health science library. *Issues in Science and Technology Librarianship* 28, http://www.istl.org/00-fall/refereed.html (accessed March 27, 2008).

Nixon, Judith M. 2000. Using a database management program as a decision support system for the cancellation of business periodicals and serials. *Journal of Business & Finance Librarianship* 5, no. 4: 3-21.

Noonan, Christine F., and Melissa K. McBurney. 2007. Application of electronic serial usage statistics in a national laboratory. In *Usage statistics of e-serials*, ed. David C. Fowler, 151-60. Binghamton, NY: Haworth Information Press.

Pesch, Oliver. 2006. Ensuring consistent usage statistics, part 1: Project COUNTER. *Serials Librarian* 50, no. 1/2: 147-61.

Pesch, Oliver. 2004. Usage statistics: Taking e-metrics to the next level. *Serials Librarian* 46, no. 1/2: 143-154.

Peters, Thomas A. 2002. What's the use? The value of e-resource usage statistics. *New Library World* 103, no. 1172/1173: 39-47.

Ralston, Rick. 2007. Assessing online use: Are statistics from web-based online journal lists representative? *Journal of Electronic Resources in Medical Libraries* 4, no. 1/2: 51-64.

Townley, Charles T., and Leigh Murray. 1999. Use-based criteria for selecting and retaining electronic information: A case study. *Information Technology and Libraries* 18, no.1: 32-99.

Trueswell, R.L. 1969. Some behavioral patterns of library users: The 80/20 rule. *Wilson Library Bulletin* 43: 458-61.

Weintraub, Jennifer. 2003. Usage statistics at Yale University Library. *Against the Grain* 15, no. 6: 32-34.

Wood, Elizabeth H. 2006. Measuring journal usage: Add a survey to the statistics? *Journal of Electronic Resources in Medical Libraries* 3, no. 1: 57-61.

Wulff, Judith L., and Neal D. Nixon. 2004. Quality markers and use of electronic journals in an academic health sciences library. *Journal of the Medical Library Association* 92, no. 3: 315-22.

Bibliography Entries [[[heading]]]

Bevan, Simon, Pete Dalton, and Angela Conyers. 2005. How usage statistics can inform national negotiations and strategies. *Serials* 18, no. 2: 116-23.

Botero, Cecilia, Steven Carrico, and Michele R. Tennant. 2008. Using comparative online journal usage studies to assess the big deal. *Library Resources & Technical Services* 52, no. 2: 61-68.

Chisman, Janet K. 2007. Electronic resource usage data: Standards and possibilities. In *Usage statistics of e-serials*, ed. David C. Fowler, 275-85. Binghamton, NY: Haworth Information Press.

Davis, Philip. 2003. Why usage statistics cannot tell us everything, and why we shouldn't dare to ask. *Against the Grain* 15, no. 6: 24-26.

Fischer, Karen. 2006. eJournal interface can influence usage statistics: Implications for publishers, libraries, and Project COUNTER. *Against the Grain* 18, no. 1: 64.

Fowler, David C. 2007. *Usage statistics of e-serials*. Binghamton, NY: Haworth Information Press, 2007. Co-published simultaneously as *Serials Librarian* 53, supp. 9, 2007.

Haddow, Gaby. 2007. Level 1 COUNTER compliant vendor statistics are a reliable measure of journal usage. *Evidence Based Library & Information Practice* 2, no. 2: 84-86.

Howard, Mark. 2002. Ejournals not cancelled. *Information World Review* 176 (January): 10.

International Coalition of Library Consortia. Guidelines for statistical measures of usage of web-based information resources (update: December 2001). International Coalition of LIbrary Consortia. http://www.library.yale.edu/consortia/2001webstats.htm.

Jasper, Richard P. 2006. 21st century shell game: Cutting serials in the electronic age. *Acquisitions Librarian* 18, no. 35/36: 161-66.

Johnson, Peggy. 1997. Collection development policies and electronic information resources. In *Collection management for the 21st century: A handbook for librarians*, eds. G.E. Gorman and Ruth H. Miller, 83-104. Westport, CT: Greenwood Press.

Kent, Allen, et. al. 1979. *Use of library materials: The University of Pittsburgh study*. New York: Marcel Dekker.

Metz, Paul, and John Cosgriff. 2000. Building a comprehensive serials decision database at Virginia Tech. *College and Research Libraries* 61, no. 4: 324-34.

National Information Standards Organization. NISO Z39.7-2002 draft standard for trial use. National Information Standards Organization. http://www.niso.org/emetrics/current/complete.html.

Pesch, Oliver. 2007. Usage statistics: About COUNTER and SUSHI. *Information Services & Use* 27, no. 4: 207-13.

Peters, Thomas A. 2001. What's the big deal? *Journal of Academic Librarianship* 27, no. 4: 302-04.

Rhine, Leonard. 1996. The development of a journal evaluation database using Microsoft Access. *Serials Review* 22, no. 4: 27-34.

Rous, Bernard, ed. 2004. *Online usage statistics: A publisher's guide*. New York: Association of American Publishers.

Shim, W., and C.R. McClure. 2002. Improving database vendors' usage statistics reporting through collaboration between libraries and vendors. *College and Research Libraries* 63, no. 6: 499-514.

Ward, Randall K., John O. Christensen, and Elizabeth Spackman. 2006. A systematic approach for evaluating and upgrading academic science journal collections. *Serials Review* 32, no. 1: 4-16.