

Statistics for electronic resources

Caroline Elizabeth Dean¹

Electronic Resources Librarian, University of Cape Town Libraries
caroline.dean@uct.ac.za

Karin de Jager²

Department of Information and Library Studies, University of Cape Town
karin.dejager@uct.ac.za

Received: 12 November 2008

Accepted: 14 April 2009

Electronic resources represent a large portion of many libraries' information resources. This paper sets out how and why international and South African librarians keep statistics for electronic resources, which statistics are kept, and what the issues and concerns are with regard to statistics for electronic resources. The responses were very similar. The concerns raised by both international and South African libraries were found to be about the continued lack of standardisation among publishers' reporting of statistics; the time-consuming nature of data collection; the reliability of usage data; the fact that data should be looked at in context; and the management of the data. A concern raised in South Africa but not in the international literature is that some librarians do not understand the basic concepts of electronic resources usage statistics.

Keywords: Electronic resources, library measurement statistics

1 Introduction

Statistics have long been used by libraries to measure the extent of their services. The usage of print collections has mainly been tracked with circulation figures and records have been kept of expenditure on resources and the number of volumes on the shelves. Libraries keep a count of the number of people entering their doors and the number of reference queries that librarians have answered.

Libraries no longer operate in a purely physical, but more often in a hybrid environment with print and electronic information resources available side by side. With the development of the Internet and the wealth of electronic resources, large portions of libraries' collections are no longer kept on shelves, but are available electronically on CD-ROM (compact disc, read-only memory), DVD (digital video disc) or the Internet. In order to report fully on their usage, libraries have to keep statistics for electronic resources as well.

Since the introduction of electronic resources in libraries, it has been generally recognised that it should be possible to obtain usage statistics from the computer systems that facilitate access to the information resources (Cox, 2003/2004: 92; Sack, 2003/2004:36). Librarians expected that this data would enable them to see how library material was being used more easily than had ever been possible in the print environment (Gallagher, Bauer & Dollar, 2005:172). Unfortunately, in the very early days of usage statistics for electronic resources, librarians were disappointed, primarily as publishers did not obtain and report usage data in a standardised way. This meant that librarians were not clear about what the usage data actually meant, and comparisons between resources were not possible as usage was being measured differently from resource to resource.

The aim of this research project was to establish which statistics libraries internationally were keeping for electronic resources and to establish where South African libraries stand in relation to the international situation. This paper therefore attempts to consider the purpose of keeping statistics for electronic resources; which statistics libraries keep and what the issues and concerns are with regard to gathering statistics for electronic resources. As in the print environment, three main types of statistics are kept for electronic resources: usage statistics, numbers of resources and expenditure.

Usage statistics consist of information such as the numbers of full-text articles downloaded from an electronic journal or a database collection of electronic journals; or numbers of searches performed on a database or platform (a platform is a website that hosts several databases); or the number of sessions, or successful connections, to a database or platform. Also included are turnaways, or rejected sessions, for an electronic journal or database.

1. Caroline Dean, Electronic Resources Librarian at the University of Cape Town Libraries

2. Karin de Jager, PhD, Associate Professor, Department of Information and Library Studies, University of Cape Town

Number of resources refers to the number of electronic resources held by a library. This information may be broken down by format, for example number of databases, or electronic journals, or electronic books. These numbers could be further broken down by subject or subscribing department. In addition to the subscribed resources, a library might keep track of the number of free electronic resources to which they provide links from their library home page or list of electronic journals.

Cost-based information such as total amount spent on electronic resources or the percentage of budget committed to electronic resources comprise the *expenditure* statistics. This information may be broken down by format, subject, or subscribing department.

Librarians can access statistical data for electronic resources from a variety of sources. Usage data is often made available to librarians through the administration functionality of each resource on the publisher's website. The librarian is provided with a username and password to log into the administration features and access the usage data. Alternatively, some publishers send files of usage data to librarians via e-mail. At present, not all publishers supply their usage data to libraries in the standardised format according to the international code of practice known as Project COUNTER, which will be discussed below.

Although a time-consuming exercise, usage statistics for electronic resources are relatively easy to gather from websites. They can demonstrate a level of actual use that was never determinable from the circulation statistics gathered from libraries' integrated systems, as one can often see the use of specific journal titles within large databases (Blake & Schleper, 2004:461). In addition, electronic resource management tools such as Open URL resolvers (for example SFX) and alphabetical electronic journal list providers (for example, SerialsSolutions, EBSCO AtoZ, TDNet Journal Manager) can provide aspects of usage statistics for electronic resources.

SerialsSolutions, an electronic journal list provider, can provide data on the number of times each electronic journal is selected from the list, a so-called "click-through". Although this information provides an indication of the level of use being made of each electronic journal, it does not give the complete picture, as it does not count direct access to journal websites independently from SerialsSolutions. SFX, an open-URL resolver, can also provide such "click-through" usage data (open-URL resolvers are systems that facilitate linking between resources (Emery, 2005:139), typically from an indexing database to a full-text electronic journal).

These tools are also able to supply data on the number of electronic journals available to library users. This is particularly useful when it comes to establishing the number of titles accessible through aggregated databases. (These are databases of mainly full-text journal content from a variety of publishers, assembled into a database by a third party publisher. Frequently, there is an embargo on the latest issues of journals in these databases.)

Furthermore, the acquisitions/serials module of the ILS (or equivalent system) can provide data on the number of subscriptions held and the expenditure made on these subscriptions. Institutional Information Technology (IT) departments keep an automatic log, or record, of websites accessed by Internet users at the institution. Libraries can estimate the number of sessions for a resource by looking at institutional web server logs; however, this data cannot supply information on searches and article downloads (Blečić, Fiscella & Wiberley, 2001: 435). Another shortcoming of using web server log data to count database usage is that if users bypass the front screen of the database, and link through to an individual journal within the database, the session would not be counted towards that database's use as the web server log would show access being made to a different website or URL (Coombs, 2005: 601).

Finally, where there is no automated alternative, librarians have resorted to counting use manually. This is an unreliable method of counting as it relies on librarians being able to watch what is happening at the resource and manually counting as each user sits down to use it. A manual session count would be used to tally the use made of stand-alone databases, particularly CD-ROM/DVD resources, physically in the libraries.

2 Standardising e-metrics

The term *e-metrics* refers to the measurement of "the activity and use of networked information" (White & Kamal, 2006: 5). The ICOLC *Guidelines for statistical measures of usage of web-based information resources* have proved to be one of the most influential outcomes from the various e-metrics projects that have been undertaken in the library environment (White & Kamal, 2006:24).

The International Coalition of Library Consortia (ICOLC) was formed in 1996. The Coalition is a group of 200 library consortia from all over the world who get together to address issues of mutual interest. In November 1998 the Group issued the first version of their *Guidelines*. The purpose of the *Guidelines* was to provide the Group members with the required information to ensure that all consortium members were receiving usage data for their licensed resources. The *Guidelines* were also meant to provide publishers with information that would enable them to understand their customers' needs. The *Guidelines* were revised in December 2001 and then again in September 2006 (ICOLC, 2006).

They are available at <http://www.library.yale.edu/consortia/webstats06.htm> [2008, July 21]. ICOLC's motivation in compiling its *Guidelines* was to meet the interests of its members "by defining and creating a common set of basic use information requirements that are an integral and necessary part of any electronic product offering".

The ICOLC *Guidelines* provide **minimum requirements** of "data elements" that have to be made available according to database, institutional IP addresses, by consortium and by time period. These are the numbers of sessions (logins), of queries (searches), of menu selections, of "full-content units examined, downloaded, or otherwise supplied", and of turnaways. The ICOLC *Guidelines* require providers to *define* each of the data elements they supply in specific *report formats*. The *Guidelines* also provide sample reports as illustrations for providers of the minimum requirements (ICOLC, 2006).

Another result of librarians' stated need for uniformity in terms of usage statistics may be seen in the development of Project COUNTER (Counting Online Usage of NeTworked Electronic Resources). It began, in 2002 (White & Kamal, 2006:22), as an initiative of the PALS (Publisher and Librarian Solutions) group in the United Kingdom and is now the international standard (Pesch, 2004a:4; Shepherd, 2006:144) for usage statistics for electronic resources. COUNTER is owned by its members who are "publishers, intermediaries, libraries, consortia and industry organisations" (Shepherd, 2006:151).

The Project's objective was to "develop agreed international Codes of Practices governing the recording and exchange of online usage data for different categories of content." Codes of Practice have been released for journals and databases (release 1 in January 2003; release 2 in April 2005; release 3 published as a draft in March 2008) and books and reference works (in 2006) (Shepherd, 2006:143). COUNTER specifies how and when usage should be counted, thereby providing consistency across publishers (Pesch, 2006:150).

The full text versions of the Codes of Practice are freely available on the COUNTER web site (www.projectcounter.org). Whereas the standards created by NISO and ISO are formal, the COUNTER Code of Practice is a "voluntary set of guidelines that address terminology, layout and format of the report, processing of usage data, what categories or filters should be available, and delivery of reports" (Pesch, 2006:148). In order to become "COUNTER-compliant" publishers must supply their customers with the usage reports that are applicable to their product (COUNTER, 2005a:9). The *Code of Practice* clearly defines all the data elements that are required in each report and provides an example of each report demonstrating the layout of the report.

Shepherd, the director of the COUNTER project, has provided a useful outline of the reasons why librarians need usage statistics:

- To "assess the value of different online products/services";
- To "make better-informed purchasing decisions";
- To "plan infrastructure and allocation of resources"; and
- To "support internal marketing and promotion of library services" (Shepherd, 2006:142).

3 Internationally recognised statistics for electronic resources

In an article written in 1999, and therefore quite dated in terms of developments around electronic resource usage statistics, Hiott discusses usage statistics at the Houston Public Library. Statistics that were kept at the time were: sessions, logins, searches, documents downloaded, and turnaways (Hiott, 1999:44). Over the years these are still the core statistics that librarians are keeping, and will be discussed in greater detail below. COUNTER defines a Session as follows:

A successful request of an online service. It is one cycle of user activities that typically starts when a user connects to the service or database and ends by terminating activity that is either explicit (by leaving the service through exit or logout) or implicit (timeout due to user inactivity) (COUNTER, 2005b).

Although the ICOLC *Guidelines* imply that the terms *sessions* and *logins* can be used interchangeably (Blecic, Fiscella & Wiberley, 2007:27; ICOLC, 2006), it is clear from the COUNTER definition that a session is far more than just logging into a database, and usually involves some activity once logged in. *Cost-per-session* figures can then be calculated by dividing the subscription cost by the number of sessions. This information allows the librarian to put a measure of value to a resource (Conyers, 2004:151; Moen, Oguz & McClure, 2004).

Commercially produced alphabetic lists of libraries' electronic journal holdings are another source of usage statistics. EBSCO's A-to-Z administration system is able to produce session, search and link-out reports that can be run according to various time frames. Session reports carry information on the number of sessions opened, including details of "average session length, total number of pages viewed, and average number of pages viewed per session" (Lingle, 2005:52).

The number of *searches* conducted can indicate the level of use being made of the resource and can be used to compare the importance of similar resources. Pesch (2004b:146) suggests that the number of searches conducted could

be equated to reference questions answered by librarians. *Cost-per-search* figures can then be calculated by dividing the subscription cost by the number of searches conducted. This information allows the librarian to put a measure of value to a resource (Conyers, 2004:151; Moen, Oguz & McClure, 2004).

The number of full-text articles downloaded is used by libraries to demonstrate the level of use of resources. Plotting the number of full-text downloads over time can illustrate trends of use in the library. Combining the number of full-text articles with the subscription cost can provide information on the cost per download, and this can be used to measure the value of the information accessed (Bevan, Dalton & Conyers, 2005:117-121; Moen, Oguz & McClure, 2004).

Turnaways (Bertot, *et al.*, 2004:31) or rejected sessions can be used to identify resources that need to have their subscriptions upgraded to allow for an increase in the number of simultaneous users for those resources (Pesch, 2004b:145). Libraries might begin their subscription to a resource with a small number of simultaneous users in order to gauge the level of real interest at the institution. A high number of turnaways indicates a real need for access to the resource and the subscribing library should consider spending more on their subscription in order to increase the number of simultaneous users.

Some libraries are interested in reporting usage of electronic resources by *location*. This might be usage by a branch library (Hiott, 1999:45), or usage by remote access. Electronic usage reported by a branch library would enable the branch librarians to identify which resources might be in need of promotion at the branch, if usage is low. Usage reported by branch also allows the branch librarians to study the pattern of use for electronic resources at the branch and will allow them to plan their services accordingly.

Information on the *numbers of electronic library holdings* in the various formats allows for the benchmarking of services and collections (Conyers, 2004:149; Pesch, 2004b:144). Counting resources by subject can identify areas that need collection development and help with collection sharing. Knowing the numbers of journals in various collections or databases can allow the librarian to perform overlap analysis to ensure that there is a minimum of duplication in purchased resources (Pesch, 2004b:144).

Expenditure on electronic resources, broken down according to databases, electronic journals and electronic books (Bertot, *et al.*, 2004:31) can allow for benchmarking against peer institutions. A “*virtual visit*” is a “hit on the library website” (Conyers, 2004:152). Hits need to be counted and added to actual visits counted by the physical library (Bertot, *et al.*, 2004: 32) to give a true reflection of the use being made of the library services.

4 Issues and concerns

A number of issues and concerns with regard to statistics for electronic resources have been identified from a survey of the international literature. The lack of standardisation was an issue back in 1999 already (Hiott, 1999:47) and is only gradually being solved by the growing adoption of COUNTER compliancy by publishers. Standardisation is not yet fully realised, with some publishers/vendors being COUNTER compliant and others not (Ferguson, 2003:32; Schmidt, 2006). When there is such a lack of uniformity, comparisons between resources cannot be undertaken (Bauer, 2001:37; Franco, 2005).

Problems also arise with interpreting usage data. This can be due to the way the information is collected and the lack of standardised terminology in the area (Franco, 2005). The basic terms need to be clearly defined (Peters, 2002:44). The COUNTER Codes of Practice contain precise definitions of all applicable terms and their guidelines on how publishers should collect and report usage data, and ensure that standardisation in this area is increasing.

The collection of usage statistics is a time-consuming process (Ferguson, 2003:32), as it typically involves going to the publisher’s website, logging in to the administration system, locating the files of relevant data, downloading them to a local computer and collating them into a programme such as Microsoft Excel. Because it is such a time consuming process it has to be asked whether “the value of the measure’s results [are] worth the effort necessary to obtain them” (Bertot, McClure & Ryan, 2001:51). Librarians want to be assured that what they are counting is worthwhile (Conyers, 2004:149) and that their statistics may be put to meaningful use. A further problem is that there is usually at least a month’s delay in statistics becoming available. Data can also be corrupted or missing (Blake & Schleper, 2004:461; Ferguson, 2003:30). These issues affect librarians’ confidence in the reported usage data. Concern has been raised whether publishers can be trusted not to inflate usage reports in order to ensure continued subscriptions (Schmidt, 2006). The independent auditing that is undertaken as part of COUNTER compliancy should address this concern.

There is also concern about how metasearching might impact on the usage statistics for individual electronic resources (Schmidt, 2006). A metasearch, or federated search, system allows the user to search across a range of resources that need not be on the same platform (Emery, 2005:139). The numbers of searches recorded for resources that are metasearchable could be inflated by users who do not discriminate between databases when selecting which to

cross-search (Stubblings & Hamblin, 2004:29), but simply select all those on offer. The usage figures will then no longer reflect someone's conscious choice to use a specific resource to meet a specific need.

There is great concern that libraries might make cancellation decisions based purely on cost and usage statistics. "... [W]e need to be careful about any inferences we make from an analysis of usage data about the needs, interests, and preferences of users" (Peters, 2002:4). The caution is that the usage data must be looked at in context (Franklin, 2005:245; Kraemer, 2006:164; Schmidt, 2006). For example, "one should not focus on raw numbers, but use compared to similar journals" (Nisonger, 2000:300).

The "access infrastructure" (off-campus access; OpenURL resolvers, metasearch systems) has an effect on usage; those journals with sophisticated linking functionality tend to show higher use (Kraemer, 2006: 169). The prominence a resource is given on a library webpage will influence how easily it is seen by users, thereby increasing its usage (Ferguson, 2003: 33). Sometimes a library will continue to subscribe to an expensive resource with high cost-per-use figures because it is an important resource for a specific department (Franklin, 2005:245; Stubblings & Hamblin, 2004:25).

Complications might arise in reporting the number of electronic resources a library holds. When it comes to counting databases, what constitutes a database? Does a collection of full-text journals constitute a database? For reporting on the number of electronic *journals* a library holds, it is often difficult to keep track of the journal titles and numbers of journal titles included in big deals, publisher collections and aggregated databases. The journal titles in these collections count towards the number of electronic journals held by the library even though the library does not actually subscribe to the individual titles. Another complication is whether a library counts a print plus electronic subscription as a print journal subscription or as an electronic journal subscription, or as both. This question is particularly pertinent when calculating expenditure on electronic resources (Conyers, 2004:150). The items that are to be counted clearly need to be well defined.

5 Investigating the situation in South Africa

The Coalition of South African Library Consortia (COSALC) was established in July 1999 (South African National Library and Information Consortium, 2007) in order to unify the efforts of the five South African higher education consortia that had been established between 1992 and 1998 (Thomas, 2007:82,83). During 2005 and 2006 COSALC underwent a restructuring process and in 2007 the body was renamed the South African National Library and Information Consortium, SANLiC (COSALC, 2007) to broaden the focus and to become more accessible to membership beyond academic and research consortia (Thomas, 2007:83).

The main project of the former COSALC was SASLI, the South African Site Licensing Initiative. At its Annual General Meeting in November 2006 it was agreed, amongst other things, that the SASLI name would fall away to be replaced by SANLiC-SL (SANLiC Site Licensing) (COSALC, 2007). In 2006 SASLI was instrumental in organising two learning opportunities for librarians on electronic resources usage statistics as a need for this had been expressed by the regional consortia. There was a session on statistics at the two-day seminar held at the CSIR Conference Centre, Pretoria, in July. The title of the event was *An integrated approach to e-content: structure through technologies and standards*. At around the same time Peter Shepherd, Director of Project COUNTER, visited the South African regional library consortia to inform librarians about COUNTER.

During 2006 it was therefore decided to investigate whether South African libraries keep and use the same statistics for electronic resources as their international colleagues. A questionnaire was designed and based on the issues that librarians face regarding electronic resource statistics internationally. On 7 November 2006 this questionnaire was e-mailed to the 23 South African academic institutions that form SANLiC. A reminder e-mail was sent on 3 January 2007 to those institutions that had not responded to the original message. On 7 August 2007 the questionnaire was again sent out to non-respondents. In the end a total of 15 responses was received, thus yielding a response rate of 65%. It is acknowledged that the number of responses is small, but as the instrument was sent to the whole population, i.e. all the South African SANLiC institutions that are academic libraries, and responses were received from across the range of institutions, it does give a reasonably accurate picture of the South African academic library situation. Respondents were asked to give their names and institutions for administrative purposes and were also asked to indicate their staff position. Confidentiality was assured and no individual institution would be identifiable from the discussion of the findings, which are summarised below.

5.1 Statistics for databases

The first question enquired which publisher or vendor produced usage statistics were kept by the individual institutions, to ascertain whether institutions are recording the database statistics that COUNTER requires from database publishers. Sub-questions referred to data that are stipulated in the COUNTER code of practice for journals and databases and are reported in the COUNTER Database Reports 1, 2 and 3. Two institutions did not answer the detail of this question, other

than to say they kept statistics only for what they regarded as the most important databases. The analysis of the question is therefore based on responses from 13 institutions. The results were as follows:

QUESTION 1: Statistics kept for databases:

- | | |
|---------------------------------|--------------|
| a) Number of searches | 13 YES |
| b) Number of sessions | 8 YES; 5 NO |
| c) Number of fulltext downloads | 12 YES; 1 NO |
| d) Number of turnaways | 11 YES; 2 NO |

Respondents were also asked to note whether they kept any additional statistics that are not required for COUNTER compliancy. Comments showed that one institution noted the numbers of PDF and of HTML full-text downloads, thus providing interesting additional information. Although the COUNTER reports for databases do not require publishers to distinguish between PDF full-text article downloads and full-text articles downloaded in HTML, this distinction is present in the COUNTER Journal Reports, specifically Journal Report 1. The databases that are basically large collections of full-text journals, sometimes report the PDF and HTML downloads at a database level as well as at the journal level. One institution noted that the number of e-mail requests was also recorded, which is not a figure that many databases provide. However, where it is provided, it gives an extra indication of how users are interacting with the database. In answering this question, a further institution commented that although they record all the statistics they only make use of the figure for the number of searches.

5.2 Statistics for electronic journals

Question 2 referred to electronic journals, whereas Question 1 had focused on databases. The question asked which publisher or vendor produced user statistics are kept by each individual institution. Electronic journals could be journals that are accessed as part of a database or collection of electronic journals, or they could be individually accessed. In both scenarios the statistics discussed in Question 2 deal with the individual journal titles. Two institutions did not answer the detail of this question, therefore the analysis for this question is done on responses from 13 institutions.

QUESTION 2: Statistics kept for Electronic Journals

- | | |
|---------------------------------|--------------|
| a) Number of searches | 11 YES; 2 NO |
| b) Number of sessions | 4 YES; 9 NO |
| c) Number of fulltext downloads | 11 YES; 2 NO |
| d) Number of turnaways | 5 YES; 8 NO |

Responses show that fewer institutions record usage data for electronic journals than for databases. Usage is recorded according to what is available from the publisher platform. For the most part the availability of data matches the information available for databases, except that the figures are recorded per journal title. In response to the last question in this section which asked about any additional electronic statistics being kept, two institutions mentioned counting usage from their electronic journal lists. Most of the commercial providers of alphabetical lists of electronic journals provide usage statistics for the list. Although COUNTER specifies that the responsibility for providing usage data lies with the publisher in this situation, the usage information that is available through these providers (for example SerialsSolutions and EBSCO) can provide a picture of how library users are accessing the electronic journals.

5.3 Statistics for electronic books

Question Three was to establish whether libraries keep statistics for electronic books to the same extent as for databases and electronic journals. All 15 respondents answered Question Three.

QUESTION 3: Statistics kept for Electronic books

- | | |
|--|--------------|
| Do you keep statistics for electronic books? | 4 YES; 11 NO |
|--|--------------|

Only four respondents indicated that they record statistics for electronic books. This small number is indicative of the slow uptake of electronic books within academic institutions in South Africa. Although more institutions are showing an interest in electronic books, many institutions are moving into this arena with caution. A recently published paper on the

cataloguing of electronic books in South Africa concluded that a substantial number of librarians professed a “lack of experience and confidence” in cataloguing electronic books (De Jager, 2007:52). It is suggested that this inexperience might also be extended to a limited ability in collecting statistics for electronic books.

5.4 Other electronic resource statistics

Question Four addressed statistical data that were not usage data, including information on the numbers of the various types of electronic resources and expenditures. The question enquired which other figures were being kept and answers were provided by 13 institutions.

QUESTION 4: Other statistics

a) Number of database subscriptions	12 YES; 1 NO
b) Number of electronic journal subscriptions	11 YES; 2 NO
c) Number of electronic books purchased	8 YES; 5 NO
d) Expenditure on electronic resources	11 YES; 2 NO

Most institutions have access to records of the number of electronic resources to which they subscribe, and they can be differentiated by format, namely database, electronic journal and electronic book. Most institutions can also access data on the expenditure figures for these resources.

5.5 Staffing

Question 5 addressed the issue of staff involved in the process of collecting usage statistics for electronic resources. Ten institutions each have one member of staff involved in collecting electronic resources statistics, while five have two staff members involved. Two institutions reported that para-professional staff do the collating of statistics, with a librarian doing the analysis and interpretation. One institution reported that their IT personnel deal with usage statistics. At 12 institutions professional staff attend to the whole task. A sub-question asked how much time staff members spent on collecting and managing statistics for electronic resources. Libraries reported spending between 30 minutes and 5 days a month on collecting and collating usage statistics. As Figure 1 illustrates, most of the institutions spend approximately 14 hours per month collecting statistics.

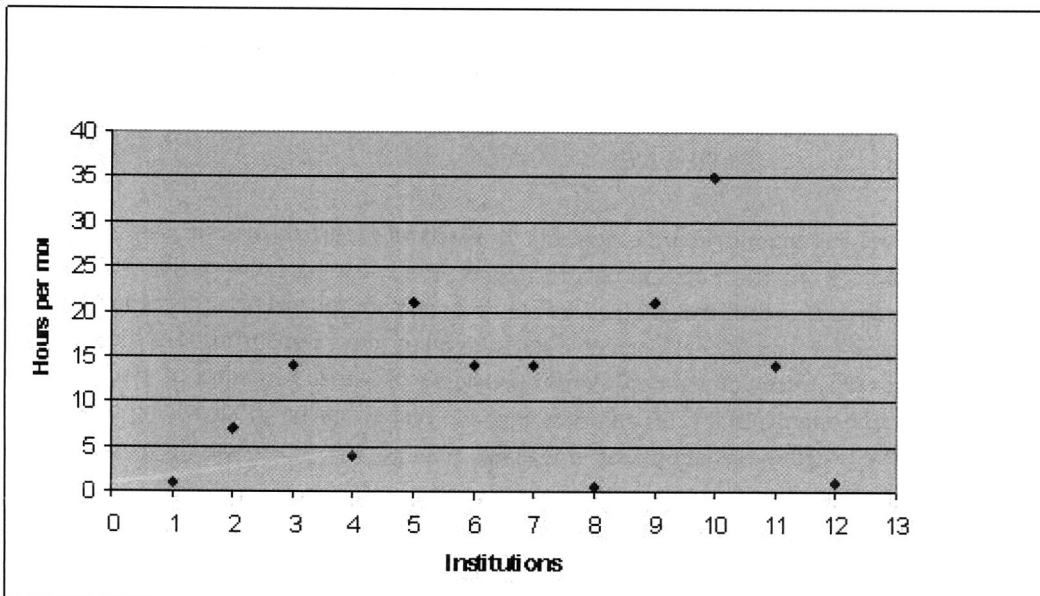


Figure 1: Time spent collecting statistics

The answers to a further sub-question that asked how usage statistics are recorded and managed, indicate that some libraries use very basic methods, while others have well-developed procedures. One institution did not answer this question at all, and two gave answers that were unusable. Of the twelve institutions that gave meaningful answers, ten indicated that they used a spreadsheet to manage their statistics. Two institutions reported that they printed out the statistics from the vendor sites and filed them.

When using a spreadsheet, institutions record the statistics by month. Four institutions reported that the statistics are made easily accessible to library colleagues on intranets or common network drives. Three institutions reported that they distribute statistics to key stakeholders in the institution.

5.6 Use of statistics for electronic resources

Question 6 focused on how libraries were using electronic resource statistics. Although all 15 institutions indicated that they used the statistics for some purpose, two did not answer the more specific sub-questions.

QUESTION 6: Do you use statistics on electronic resources? Yes or No?	15 YES
If YES, do you use statistics to:	
a) demonstrate level of use?	11 YES; 2 NO
b) demonstrate costs, eg. cost per fulltext download, cost per session, cost per search?	12 YES; 1 NO
c) determine correct subscription level in terms of number of simultaneous users?	10 YES; 3 NO
d) monitor number of turnaways?	10 YES; 3 NO

It was noted above that low usage statistics or decreasing usage statistics can be used as indicators that training and marketing are required for certain resources. These same low usage figures might also be used as a reason not to renew subscriptions, although when used for this purpose the statistics must be looked at in context and not in isolation.

5.7 COUNTER awareness

Question 7 enquired whether respondents were aware of the Project COUNTER initiative and asked respondents to explain how important it was for their purposes that usage statistics are COUNTER compliant. Out of 15 institutions, two responded that they were not aware of the COUNTER initiative. Although they had heard of the initiative, another two institutions did not know anything about it. The 11 institutions that knew about Project COUNTER counted it as important as it imposed standardisation on the usage statistics produced by publishers. Given the fact that Peter Shepherd (Director, COUNTER) did a series of presentations during July 2006 at the various centres throughout South Africa and at the SASLI seminar in Pretoria, it is disappointing that not all institutions and librarians are fully versed in Project COUNTER and the advantages it brings to the world of electronic resource usage statistics.

5.8 Other issues

The last open question allowed respondents to raise any additional issues that had not previously been mentioned. Respondents noted that:

- There are still publishers who cannot provide libraries with usage statistics. This is particularly relevant in the South African situation, with major publishers e.g. LexisNexis and Jutastat unable to supply this information.
- Libraries are considering developing their statistics processes. One institution reports looking into subscribing to ScholarlyStats to help manage their statistics. Two respondents already subscribe to this service. One institution is looking for ways to determine usage by faculty so that they will be able to target specific areas that need training or for the marketing of resources.
- Three respondents asked whether the effort entailed in collecting statistics made it worthwhile to collect statistics for all resources, or whether the effort should be focused on a selection of key subscriptions only. Although collecting limited statistics might sound like a way to manage the amount of time spent on collecting usage statistics, it does mean that the library will be unable to accurately report the level of use being made of their electronic resources.
- Although strides have been made in terms of usage statistics for online resources, this information for networked CD-ROMs is lacking. This concern was raised by two respondents. As already discussed, libraries resort to keeping manual statistics for CD-ROM databases, whether stand-alone or networked. It is highly probable that this manual count does not give an accurate portrayal of the use of CD-ROM databases. These usage figures are also not comparable against usage figures for other resources.
- When libraries are short staffed, collecting usage statistics becomes very low priority. When it appears that little use is being made of usage data, busy staff will tend to place the task of collecting the information low on their priority list.
- Some publishers have very user unfriendly systems for their statistics.
- A further issue became apparent from the way in which the survey questionnaire was completed. A few institutions completed the survey in a very rudimentary fashion. While this might be indicative of an unwillingness to complete survey questionnaires, or a sign of not having the time to do so, it could also suggest that the respondents did not have a good understanding of the collection and use of statistics for electronic resources. This could be an expression of a training need in this area, which is corroborated by De Jager's (2007:53) findings regarding the cataloguing of electronic books in South Africa.

6 Discussions: statistics for electronic resources in South Africa

The reasons a number of South African libraries keep statistics for electronic resources fall (with one exception) fairly neatly into Shepherd's outline of reasons in the international libraries domain (Shepherd, 2006:142).

a) To assess the value of different online products/services

As discussed above, most of the South African libraries that responded to the survey (12) report using usage statistics and cost of subscriptions to calculate cost-per-use figures, for example cost per search. These figures give the libraries a measure of the relative values of resources. This allows a certain degree of comparison to be made of resources.

b) To make better informed purchasing decisions

Eleven of the South African libraries that responded to the survey collect usage statistics to monitor the level of use being made of resources. Usage figures are being used to support renewal and cancellation decisions for subscriptions. Turnaway counts are being used to ensure that subscription levels are at optimum for the resources that have a simultaneous user limit.

c) To plan infrastructure and allocation of resources

The authors could not identify planning and allocation of resources as a reason for keeping statistics from the questionnaire responses from South African libraries.

d) To support internal marketing and promotion of library services

Two South African libraries were found to use low usage figures to identify resources that require promotion to users. The circulation of usage reports to key stakeholders in institutions accomplishes a degree of marketing of library services. Where available, some South African libraries report that they are using usage statistics for specific resources to make comparisons with their peer institutions.

Like their international counterparts, a number of South African libraries report that they keep the following statistics for electronic resources:

- Searches
- Sessions
- Full-text downloads, including the distinction between PDF and HTML where provided
- Turnaways
- Number of subscriptions
- Expenditure
- Other statistics provided by publishers that are not COUNTER compliant: hits; visits; e-mails sent from database

No South African library reported keeping statistics on virtual visits.

7 Conclusion and recommendations

It is apparent from the varied responses to the questionnaire that there are gaps in some librarians' knowledge regarding the potential uses for statistical data for electronic resources. Not all librarians are aware of Project COUNTER and the work being done to standardise electronic resource usage data. Ideas on how usage data might be used to promote the library's standing in the institution also need to be disseminated more widely. One way to do this is to be able to demonstrate how well one's own library services are performing against peer institutions in the country. With coordination at the Consortium level, it might be possible to get agreement from all participants on requirements for reporting usage data, thereby setting up the framework for establishing benchmarks for electronic resources in South Africa.

Training events could be scheduled under the auspices of SANLiC to ensure that the library staff who are responsible for collecting electronic resource usage data, understand the concepts and the greater context of what they are doing. Usage data can be used not only to demonstrate trends with individual resources and to identify little-used resources for cancellation or marketing, but also to compare library performance against peer institutions in a national and international arena. For the effort associated with collecting and collating usage data, the library staff doing the job should realise how the information is used, or could be used, to promote their library.

References

- Bauer, K. 2001. Indexes as tools for measuring usage of print and electronic resources. *College and Research Libraries*, 62(1): 36-42.
- Bertot, J.C., McClure, C.R., Davis, D.M. & Ryan, J. 2004. Capture usage with e-metrics. *Library Journal*, 129(8): 30-32.
- Bertot, J.C., McClure, C.R. & Ryan, J. 2001. *Statistics and performance measures for public library networked services*. Chicago: American Library Association.

- Bevan, S., Dalton, P. & Conyers, A. 2005. How usage statistics can inform national negotiations and strategies. *Serials*, 18(2): 116-123.
- Blake, J.C. & Schleper, S.P. 2004. From data to decisions: using surveys and statistics to make collection management decisions. *Library Collections, Acquisitions and Technical Services*, 28: 460-464.
- Blecic, D.D., Fiscella, J.B. & Wiberley, S.E. 2007. Measurement of use of electronic resources: advances in use statistics and innovations in resource functionality. *College and Research Libraries*, 68(1): 26-44.
- Blecic, D.D., Fiscella, J.B. & Wiberley, S.E. 2001. The measurement of use of web-based information resources: an early look at vendor-supplied data. *College and Research Libraries*, 62(5): 434-453.
- Conyers, A. 2004. E-measures: developing statistical measures for electronic information services. *VINE: The Journal of Information and Knowledge Management Systems*, 34(4): 148-153.
- Coombs, K. 2005. Lessons learned from analyzing library database usage data. *Library Hi Tech*, 23(4): 598-609.
- COSALC. 2007. General Communiqué, March 2007. [Online]. Available: <http://www.cosalc.ac.za/communique.html> [2007, December 4]
- COUNTER. 2005a. The COUNTER Code of Practice for Journals and Databases. Release 2. [Online]. http://www.projectcounter.org/r2/COUNTER_COP_Release_2.pdf [2008, January 29].
- COUNTER. 2005b. The COUNTER Code of Practice for Journals and Databases. Release 2. Appendix A. Glossary of terms. [Online]. http://www.projectcounter.org/r2/Appendix_A_Glossary.pdf [2008, January 17].
- Cox, J. 2003/2004. Don't confuse price with value – in academic publishing, electronic is better. *Against the Grain*, 15(6): 92-93.
- De Jager, K. 2007. Opening the library catalogue up to the web: a view from South Africa. *Information Development*, 23(1): 48-53.
- Emery, J. 2005. Beginning to see the light: developing a discourse for electronic resource management. *The Serials Librarian*, 47(4): 137-147.
- Ferguson, A.W. 2003. Digital information use statistics: useful? Useless?. In *Charleston Conference Proceedings 2002*. R. Bazirjian & V., Eds. Westport, Conn: Libraries Unlimited. 27-35.
- Franco, T. 2005. The evaluation of electronic resources as a strategic factor in the decision making process: tools, critical points, feasible solutions. E-LIS. [Online]. Available: <http://eprints.rclis.org/archives/00004163/> [2008, February 7].
- Franklin, B. 2005. Managing the electronic collection with cost per use data. *IFLA Journal*, 31(3): 241-248.
- Gallagher, J., Bauer, K. & Dollar, D.M. 2005. Evidence-based librarianship: utilizing data from all available sources to make judicious print cancellation decisions. *Library Collections, Acquisitions and Technical Services*, 29:169-179.
- Hiott, J. 1999. Making online use count. *Library Journal*, 124(16): 44-47.
- ICOLC. 2006. Revised guidelines for statistical measures of usage of web-based information resources [Online]. Available: <http://www.library.yale.edu/consortia/webstats06.htm> [2008, January 28].
- Kraemer, A. 2006. Ensuring consistent usage statistics, Part 2: working with use data for electronic journals. *The Serials Librarian*, 50(1/2): 163-172.
- Lingle, V.A. 2005. Implementing EBSCO's A-to-Z and LinkSource products for improved electronic journal management, In *Electronic journal management systems: experiences from the field*. G. Ives, Ed. Haworth Press, 2005: 43-54.
- Moen, W.E, Oguz, F. & McClure, C.R. 2004. The challenges of nonstandardized vendor usage data in a statewide metasearch environment: the library of Texas experience. *Library Quarterly*, 74(4): 403(20). [Online] Available: <http://infotrac.galegroup.com> [2008, February 7].
- Nisonger, T.E. 2000. Usage statistics for the evaluation of electronic resources. *Library Collections, Acquisitions and Technical Services*, 24: 299-302.
- Pesch, O. 2004b. Usage statistics: Taking e-metrics to the next level. *The Serials Librarian*, 46(1/2):143-154.
- Pesch, O. 2004a. What is COUNTER and what does it mean to you? *At Your Service: An EBSCO Publication for Serials Librarians and Information Professionals*, (68): 4-6.
- Pesch, O. 2006. Ensuring consistent usage statistics, Part 1: Project COUNTER. *The Serials Librarian*, 50(1/2): 147-161.
- Peters, T.A. 2002. What's the use? The value of e-resource usage statistics. *New Library World*, 103 (1172/1173): 39-47.
- Sack, J. 2003/2004. The beginning of value assessment: usage information in the e-journal age. *Against the Grain*, 15(6): 36-40.
- South African National Library and Information Consortium. [2007]. About SANLIC. [Online]. Available: <http://www.cosalc.ac.za/home.html>. [2007, December 4].
- Schmidt, J.K. 2006. CDL-Mid-Winter discussion (#2): Use statistics for electronic resources. E-mail submitted to Listserv: colldv1@usc.edu 16 February [2006, February 16].
- Shepherd, P.T. 2006. COUNTER: usage statistics for performance measurement. *Performance Measurement and Metrics*, 7(3): 142-152.
- Stubbings, R. & Hamblin, Y. 2004. How to make your e-resources earn their keep. *Serials*, 17(1): 25-30.
- Thomas, G. 2007. Academic libraries in the second decade of democracy: critical issues and challenges. In *Libraries for the future: progress and development of South African Libraries*, T. Bothma, P. Underwood and P. Ngulube, Eds. Pretoria: Library and Information Association of South Africa, pp.71-87.
- White, A. & Kamal, E.D. 2006. *E-metrics for library and information professionals*. London: Facet Publishing.