ELECTRONIC JOURNAL USAGE STATISTICS IN PRACTICE

Tony Kidd

This paper reflects on the growing importance of usage statistics, citing increased provision of electronic journals, their increasing use, and in particular their increasing cost. The need selectively to promote particular services and titles, and the place of performance indicators, are also covered. Particular examples of the value of usage statistics are given in relation to budget allocation, collection management and collection development, and marketing of journal availability to users. The importance of statistics that are comprehensive, comparable, and easy to analyse is stressed, and current efforts to increase standardisation are supported. This paper is adapted from a presentation prepared for a one-day seminar on Usage Data and E-journals, organised by the UK Serials Group in October 2001.

Tony Kidd, Head of Serials/ Document Delivery, University of Glasgow Library, Hillhead Street, Glasgow G12 8QE
Tel: 0141 330 6778
Email: t.kidd@lib.gla.ac.uk

Introduction

This paper sets out to examine some of the practicalities of electronic journal (e-journal) usage statistics. Through a consideration of the reasons why measurement is important, indeed vital, some of the purposes for which usage statistics can be used, and have been used, are described.

It is useful to go through some of the reasons why library and information staff are so concerned to have accurate and complete e-journal statistics. Such statistics are vital for others as well, especially publishers, but this paper is written from the library standpoint, in particular from the academic library. Other papers in this issue cover views from elsewhere in the information landscape.

Increasing provision of e-journals

From a standing start, the availability of e-journals has increased exponentially over the last five years or so. Virtually all peer-reviewed STM (science, technology, medicine) journals are now available in electronic format, with social science journals not far behind, and arts and humanities journals catching up fast. Again after a slow start, libraries are now providing access to more and more of these online versions, in response to demand from their users. Although e-journals are no panacea (apart from cost problems, there are still-valid concerns over archiving and access on and off campus), it seems certain that electronic provision can only continue to increase.

Table 1 gives some figures on numbers of e-journals at Glasgow University Library, indicating doubling of provision each year between 1997 and 2000. This year the growth rate has slowed

down to 50%,
although the
numerical
increase is
the same as
the previous
year. It is likely
that this
comparative

Table 1: Electronic journals at Glasgow University Library	
Year	Numbers available
1997	362
1998	740
1999	1545
2000	3018
2001	4507

slowdown will continue, although healthy
increases will still take place for a few years.
The distribution of our subscriptions in 2001 is
interesting:

- Print only   3829
- Electronic only   2641
- Print + electronic 1866

These figures show the continuing importance
of print journals, but also show that the number
of titles available online (4507) has for the first
time exceeded the number of print only
subscriptions. It should be said that Glasgow is
not unusual in this distribution, and indeed
there are other UK academic libraries that have
made a much more determined shift towards
electronic access.

Increasing use of e-journals

Publishers and libraries have been making more
e-journals available, and this availability has been
more than matched by increasing use of this
format. Table 2 shows summary results from a
survey carried out at Glasgow University in

Spring 2001, indicating average numbers per
month per academic staff member of various
types of use of periodicals.

This survey is in fact a very good indication
of the importance of usage statistics, together
with the difficulties arising from their present
provision. The survey was undertaken to provide
data to enter into the library's budget allocation
formula. This is a very complicated calculation
which includes 'use' among a number of different
variables combined to determine the distribution
of library funds for purchase of serials, books and
document delivery services, between the eleven
different faculties operating at Glasgow. For this
if for no other reason, accurate and complete
usage statistics are thus of intense practical
interest to library and academic staff – but
'accurate' and 'complete' are both highly
misleading adjectives to apply to the statistics
that are in fact available. The problems of
objectively measuring the use of print journals
are well known and will not be repeated here.
At present, it is nearly as difficult to get
comprehensive and comparable statistics of
electronic use. In theory, print and electronic
usage would then be combined to obtain overall
figures. Because of the problems of measuring
actual use, we were reduced to surveying staff
and research students, asking them to estimate
their use, with all the attendant problems of
response rates, conscious or unconscious bias in
responses, etc. As can be imagined, the results are
controversial, but nevertheless are presented here.

Table 2: Periodical usage at Glasgow University Library: Survey of academic staff, Spring 2001

Faculty Av no/month	Periodical loans	Current issue consults	Pre-current volume consults	E-journal issue accesses
Arts & Divinity	2.0	7.0	5.9	2.4
Education	1.0	3.9	3.8	1.4
Law & Financial Studies	1.5	4.8	6.1	5.7
Social Sciences	1.7	5.5	4.5	7.7
Biomedical & Life Sciences	0.9	6.1	6.6	18.7
Comp, Stats & Math Sciences	0.8	5.2	5.5	10.9
Engineering	0.4	1.9	2.2	5.9
Clinical Medicine	0.4	2.9	2.1	11.6
Dentistry	1.6	4.1	5.5	2.8
Physical Sciences	1.6	10.1	8.0	8.8
Veterinary Medicine	1.5	5.4	3.9	10.0
TOTAL	**1.1**	**5.4**	**4.9**	**10.3**

Despite the shortcomings, the results probably do illustrate at least approximately the 'correct' usage. Taking them at face value, we can see that, in broad terms, there is very roughly equal consultation of print (11.4/month) and electronic (10.3/month) formats – though this assumes, dubiously, that 'loans', 'consultations', and 'accesses' are all indicators of a similar level of use. This in itself is a remarkable statistic given the following three considerations. E-journals did not really exist five years ago. There is a the continuing preponderance of print journals at Glasgow (though the more heavily used journals are more likely to be available online). And there are still rather few online backruns, but perhaps this is a reflection of use heavily weighted towards current issues, especially for STM journals. Variations between faculties are also pronounced, mirroring different usage patterns, and different levels of e-journal availability. Biomedical and Life Sciences staff are the heaviest users of e-journals (absolutely, and comparatively, apart from Clinical Medicine), partly but only partly because 65% of their current journal subscriptions are available online at Glasgow, compared with 18% of Arts and Divinity faculty titles.

To try to overcome some of the problems of 'self-measurement' of e-journal use, we have also co-operated with our Computing Services to measure directly e-journal accesses on a faculty basis. This information will never be available to us even from perfectly compiled and comprehensive publisher statistics, because almost all of Glasgow's web traffic goes through local caches for local traffic management purposes, and is not related, as far as the publisher is concerned, to individual or departmental IP addresses. We provided the Computing Service with URLs for the majority, but of course not all, of the e-journal services used at Glasgow, whether publisher or intermediary sites. They in turn could give us figures of usage by different faculties, or at least from networked PCs located within different faculty buildings and clusters. Figures obtained by this method in fact correlated quite well with those from the user survey, showing that at least half of all e-journal use comes from the Clinical Medicine and Biomedical and Life Sciences faculties. This method of measuring usage is probably rather more objective than surveying users, but does have its own disadvantages:

- Computing Service resources do not allow continuous measurement by URL in this way, so sampling is carried out, over various different time periods;
- Although it would be theoretically possible to measure use according to an individual user's sign-on, which would reveal status (staff/postgraduate/undergraduate) and department/faculty, this is not practical. So usage figures are approximate: in particular, almost 30% of use cannot be classified by faculty, because it takes place in the University Library, or in multi-faculty clusters, etc.;
- Accesses only are being measured: there is no distinction, for example, between viewing of full-text articles, tables of contents, and just the publisher's home page;
- URLs provided to the Computing Service must be continuously updated, though this is not usually necessary at a title-by-title level, and there is a certainty that some use will not be covered;
- It is difficult to measure use of combined database/e-journal services, such as the Ovid Biomedical Service. For budget distribution it was necessary to do this at Glasgow, as the Library pays centrally for databases, and therefore their use (asymmetrically distributed by faculty) should not count as data for the formula fund distribution. In fact, ways round this particular problem were worked out, using data provided by the service itself.

Increasing cost of e-journals

Libraries find it hard to separate the cost of electronic from print subscriptions (for example, because of the provision of print + online subscriptions by a number of publishers, and the existence of bundle deals based on the current level of print subscriptions, limiting or banning cancellations). But there is no doubt that the percentage of budgets spent on electronic access is rising. The PURCEL (2000) investigation found that 16% of UK university library budgets was spent on e-access (databases and e-journals), although some would feel that this is now an underestimate.

The increasing percentage of the budget spent in this way, combined with the ever-growing

pressures on library budgets in general and the need to demonstrate value for money wherever possible, leads to the demand for more comprehensive and reliable usage statistics, in order to satisfy budget holders within the library, within the institution, and indeed at national funding council level. This information was never really available for print subscriptions, but expectations are higher on the electronic side, and there is an understandable, and justified, feeling that much more detailed and accurate information should be provided for online journals. Although raw usage levels in themselves will never be the sole justification for subscriptions – small research groups or departments require protection, indications of quality such as impact factors will remain important, large research libraries at least may still require some depth and continuity in their collections, space for new journals must be available, local collaboration is likely to play a more important role, and there will be different considerations for journals primarily catering to student use as opposed to staff use – reliable usage figures will have a very important part to play in collection management and collection development.

Increasing usage, and so far all indications are that e-journal usage does continue to increase dramatically from year to year, can also be used to argue, with real and convincing evidence, for increased resources for the library. OhioLINK is a well-known example (Sanville, 2001) where rising usage statistics (easier to analyse in that case because services are mounted on central servers within Ohio rather than obtained from publishers' websites) have been strenuously marketed, at least in part to maintain and increase state funding where possible.

It is also likely, I believe, that pricing models will evolve towards more reliance on usage. In particular, the present arrangements with a number of large publishers, where access to a bundle of all their titles, or particular subdivisions, is based on a surcharge on the value of a library's current subscriptions, with no, or very limited, allowance for cancellations, are unsustainable, given that libraries have to cope with year to year fluctuations in their budgets, which are not increasing in real terms. In the short term, cancellations can, sub-optimally, be

made of other publishers' titles, but this is not possible for ever, especially if more publishers move in this direction. As more libraries do without print subscriptions on the other hand, the concept of the 'subscription' begins to lose meaning, to be replaced, or at least supplemented, by a more seamless transition from the acquisition/subscription end of the spectrum towards access/document delivery of individual articles. For this to work, a prerequisite is usage statistics that are standardised, accepted and understood by libraries and by publishers, to allow meaningful choices by libraries. This is also necessary if devolved budgeting is in operation, and costs of bundled deals are to be shared equitably between faculties – at present, we find this an almost impossible task at Glasgow.

For this purpose, figures provided by publishers and intermediaries are indispensable. A by-product of our provision of URLs to Glasgow's Computing Service was the production of comparative figures for the use of different publishers' and intermediaries' services. These statistics, while interesting and instructive, are not reproduced here. Not only do they suffer from most of the disadvantages of the Computing Services figures outlined in the previous section, but there is the additional problem that, for example, it is impossible to know which publishers (and titles) are represented, and to what extent, in the statistics for e.g. SwetsnetNavigator or CatchWord. Particular publishers, and titles, could be included in figures for a publisher's website, plus perhaps more than one intermediary or aggregator, and it is essential that reliable figures include usage from all different channels (including use from following links) – with the channels also identified, as libraries may be paying some kind of subscription to different intermediaries. As indicated in other papers, groups are working hard on these various problems right now, seeking and finding solutions, and developing standards (PALS Usage Statistics Working Group; ICOLC, 2001). Nevertheless, even the very raw Computing Service figures give an inkling of how we might develop and use this information in the future, allowing us to compare usage of different publishers and titles, to calculate 'cost per use' for different services and titles, and to

make subscription and document delivery decisions accordingly.

Despite current limitations, we have used statistics provided by publishers for collection management purposes. An example is Johns Hopkins University Press's Project Muse service, where we have adjusted the individual online titles that we subscribe to on the basis of usage figures (over a two year plus period: usage does vary considerably over shorter timescales, and it is necessary to take a reasonably long view).

In addition to comparability, comprehensiveness is also important. Swets Blackwell's most recent survey of publishers (Swets Blackwell, 2001) suggests that around 40% provide usage statistics. The situation is slowly improving, although with some way to go, but can only be encouraged by the adoption of industry standards.

Promotion of e-journals

One reaction to low usage as revealed by statistics is to cancel the subscription. An alternative reaction is to market the journal(s) more positively, and this might well be an initial step, especially if Subject Librarians or other staff felt that a title or service 'ought' to be consulted more intensively. At Glasgow, we initially put some effort into raising the profile of e-journals in general, by means of 'electronic parties', university newsletter articles, e-journal lists on the library website, etc., but we now concentrate more on analysing usage statistics provided by publishers, etc., though still in a fairly rudimentary way, in order to identify services where we feel that usage should be encouraged.

For example, JSTOR provides very useful backfiles particularly for arts and social science journal, and now for science journals also, but we felt that it was being underused, perhaps because arts and humanities staff were less used to

e-journals in general. Table 3 provides the evidence from statistics provided by JSTOR, showing that Glasgow's usage was not dissimilar to JSTOR's 'all sites average', but was significantly lower than the averages for 'UK large' libraries. So we have arranged special sessions on JSTOR for staff and researchers from the Arts faculty, and hope that increased use will result.

A further example is Elsevier's ScienceDirect service, where we have access to all their titles going back to 1995. Usage of this service is very heavy, although perhaps not quite so heavy as might be expected from the high print+online cost of the service. Elsevier's comprehensive usage statistics also show that there is a very wide variation in the usage per title. Although there is certainly good use of some titles where Glasgow has never had a print subscription (or where the print subscription was cancelled a number of years ago), the titles with a current print subscription have been consulted four to five times as often on average as those with no current subscription, which I suppose is some vindication at least of the library's collection development policies. Around a quarter of the titles available through ScienceDirect have been consulted ten times or fewer over a year. These figures cannot really be used at present to inform cancellations, given restrictions on cancellations, but we are planning to publicise the service more widely – for example, by making it the 'Electronic Resource of the Month' – in order to encourage further use (although there is always a balancing act to be followed in cases like this, given that we might not be able to afford to continue a subscription in the future).

Performance indicators

Related to the need to be able to justify all expenditures is the growing emphasis on

Table 3: JSTOR Glasgow University Library statistics, 2000/01							
	BROWSING				**VIEWING**	**PRINTING**		
	Title	**Vol/Iss**	**TOCs**	**Citations**	**Pages**	**Jprint**	**PDF**	**PS**
Accesses (GUL)	966	5,855	5,940	1,178	13,862	61	6,275	142
Average (UK Large)	2,512	7,434	7,150	2,006	25,198	187	8,523	118
Average (All sites)	993	2,842	2,328	1,389	11,962	201	3,792	42

performance indicators. Stephen Pinfield (2001), in a recent article on current issues in managing electronic library services for higher education libraries, highlighted the development of performance indicators, but commented that 'most usage statistics from web resources have to be treated with caution especially when trying to draw comparisons across different services. This makes the production of reliable performance indicators for e-libraries a challenge.'

SCONUL (the [UK] Society of College, National & University Libraries) has an Advisory Committee on Performance Indicators (now Performance Improvement), and the SCONUL annual statistics are beginning to include measures related to e-journals. The ARL e-metrics development (Association for Research Libraries, 2001) is also wrestling with this topic, and the EU-funded EQUINOX project has addressed the 'need of all libraries to develop and use methods for measuring performance in the new networked, electronic environment'. However, until usage statistics become more standardised and comparable, it is difficult to produce reliable performance indicators in this field.

What is to be measured?

As with so many aspects of usage statistics, the important thing here is to obtain some uniformity across statistics available from different sources. The different purposes for which usage statistics are required have been outlined above, and for most of these purposes viewing of a full-text article is probably the most meaningful single statistic which should be recorded, despite the fact that this still does not indicate whether viewing (or printing) of any given article has actually been 'useful' for the member of staff or student. Article recording should be at title level, to be cumulated by publisher if required and relevant. The volume/year of the article is an important secondary requirement, primarily to allow a library to see whether demand is principally for this year's current issues, or whether there is a significant tail of demand for articles from earlier years – this could be particularly important if there are charging implications for retaining access to earlier years' output. Information on abstracts and tables of contents accessed also helps to round out the

pattern of use, but in the interests of simplicity and comparability is perhaps less essential – although in certain subjects consultation of an article abstract can be just as 'valid' as reading the article itself.

Comparison over time is essential, and cumulation by calendar month is probably the most useful standard here, though weekly figures might also be helpful. Figures for use at different times of day can be most effective in showing budget holders, etc., the importance of 'out of hours' consultations, and therefore the importance of funding electronic access which is not limited to the hours when the library is open.

If comparability was achieved, most library staff would be happy for statistics to be retained on publishers' and intermediaries' websites – comparability is essential to allow statistics for titles available from different sources to be combined. These could then be retrieved when required, and downloaded into a standard package such as Excel for further analysis at the library level. Statistics should be retained by publishers for at least a couple of years.

Summary

Usage statistics are required because of the growing importance, and cost, of e-journal provision. They can be employed for budget allocation purposes, to aid collection development (arguing for increased resources, managing the spectrum between acquisition and access), and to inform marketing efforts by library staff.

For all these purposes, it is essential that usage statistics are universal, comparable, and easy to analyse, a goal that is closer now than it has been, thanks to the initiatives underway at present.

References

Association of Research Libraries. *Measurements for electronic resources (E-metrics)* [http://www.arl.org/stats/newmeas/emetrics/index.html].

Elsevier Science. *ScienceDirect* [http://www.sciencedirect.com/].

EQUINOX: Library Performance Measurement and Quality Management System [http://equinox.dcu.ie/index.html].

ICOLC [International Coalition of Library Consortia]. *Guidelines for statistical measures of usage of web-based information resources. December 2001 revision* [http://www.library.yale.edu/consortia/2001webstats.htm].

JSTOR [http://uk.jstor.org/].

PALS Usage Statistics Working Group [http://www.usagestats.org/].

Pinfield, S. (2001) Managing electronic library services: current issues in UK higher education institutions. *Ariadne*, 29, 2001 [http://www.ariadne.ac.uk/issue29/pinfield/].

Project Muse: Scholarly Journals Online [http://muse.jhu.edu/].

PURCEL: Purchasing decisions of electronic resources in higher education institutions. *Draft report on a research study for JISC JCALT*. Sunderland, 2000 [http://www.library.sunderland.ac.uk/jisc/Final%20report.pdf].

Sanville, T.J. (2001) A method out of the madness: OhioLINK's collaborative response to the serials crisis. *Serials*, 14 (2), July 2001, 163-177.

SCONUL Advisory Committee on Performance Improvement [http://www.sconul.ac.uk/performance_ind/index.htm].

SCONUL. *Annual library statistics*. London, 2001.

Swets Blackwell's 2000 e-journal survey (2001). *Homepages*, 1 (2), Spring/Summer 2001, 10-11 [http://www.swetsblackwell.com/custmag.htm].

———