

Usage Statistics for Decision Making

Overview of an event organized by UKSG at the Institute of Materials, Minerals and Mining in London on 10 November 2010

TERRY BUCKNELL

Electronic Resources Manager, University of Liverpool

JILL TAYLOR-ROE

Head of Liaison & Academic Services, Newcastle University

SELENA KILLICK

Research & Development Officer, Cranfield University

SARAH THOMPSON

Content Acquisition Librarian, University of York

Libraries have access to more data than ever before about the way that their electronic resources are being used by readers, and advances in standardization led by the COUNTER initiative are making these more accessible and reliable as a basis for decision making. At the same time, public sector institutions are being asked to make extremely difficult choices about priorities for spending. Responding to these concerns, UKSG hosted a new seminar in November 2010 in the stately London setting of the Institute of Materials, Minerals and Mining offering a number of case studies to demonstrate how usage statistics are informing the decision-making process in real life, and highlighting some of the perils and pitfalls of taking them at face value. The Editors of *Serials* thought that it would be useful to provide brief summaries of the papers presented for those who were unable to attend.

Terry Bucknell from the University of Liverpool started the seminar with his very personal presentation, **Apples, oranges and other unfair comparisons: ensuring decisions are based on reliable information**

More and more librarians are getting into the habit of dividing subscription costs by their usage statistics to arrive at the cost per download (CPD) and using this figure to make cancellation decisions, comparing titles between publishers and between platforms. In response, Terry suggested, it appears that publishers are looking to (artificially) drive up usage.

He pointed out that COUNTER aims to provide usage statistics that are consistent, credible and compatible. They use the same definition of what and how to count; they are regularly audited and they are delivered in the same formats. But this does *not* imply that COUNTER usage statistics for different journals on different platforms are *directly*

comparable. In COUNTER reports, every download should be equal, but Terry queried whether this is really so:

- if a platform lacks abstract landing pages then article-level links from link resolvers, Google, PubMed, etc. will lead to the full-text HTML, boosting usage figures if the user only wanted the PDF
- we know that medics read far more articles than mathematicians, for example, but that mathematicians spend far longer reading each article. Is one reading of a medical article really equal to one reading of a maths article? Much of the content in journals like *Nature*, *Science* and the *BMJ* consists of brief new stories or commentary. Do these really have the same value as a research article?
- reports may contain spikes in usage. They may signify misuse (systematic downloading) or genuine heavy use, e.g. a class.

The 'cost' side of the equation is not without its pitfalls either. In the last year or two there have been large fluctuations in exchange rates which have had dramatic effects on prices, and thus CPD calculations, even if invoiced prices have been frozen.

Terry then questioned the CPD calculation itself. Is it just a crude way to normalize usage for differently priced journals? For a rigorous unit cost calculation, librarians need to include associated staffing and systems costs. Furthermore, it is the usage of that one year's content in all future years that would be needed for a true unit cost calculation (for most subscription models). In CPD calculations, most librarians approximate that for the usage of all available content in the current year. A look at COUNTER Journal Report 5 appears to prove that is not too bad an approximation to make!

Finally, Terry sounded a note of caution. Whilst it is simple to generate CPD figures, there is a danger that decision makers will take them as the gospel truth, while librarians know (some of) the uncertainties and approximations that lie behind them. Perhaps librarians need to make this explicit by expressing them as a range rather than a single figure (e.g. £1.63 ± £0.34 rather than £1.63) and should provide a commentary to accompany any usage figures that will be used to make decisions, for example highlighting journals with spikes in usage, with limited years of content online, with a high proportion of HTML downloads, or titles invoiced in currencies that have significantly changed their value in the last year.

Jill Taylor-Roe of the University of Newcastle followed with her presentation, **Patron-led acquisition of e-books: managing the risks and exploiting the benefits**

Like most academic libraries, the Library continually seeks ways to address the perennial student refrain of: "We want more books!" In January 2010 the Library made a successful bid to the University's Teaching & Learning Innovation Fund to support an experiment with the usage-led (patron-led) e-book acquisition model offered by EBL (E-Books Library). This was just in time for the National Student Satisfaction Survey (NSS)! Around 112,000 e-book records, for titles selected by library liaison staff, were loaded into the library catalogue. In order to help identify if this type of service would satisfy the schools who most

frequently complained about a lack of books, EBL were persuaded to insert a drop-down menu which required all users to identify themselves when they logged on to use EBL. Initially, a limit of three requests per user per day was set, with additional or expensive requests mediated by the liaison librarian. Titles were purchased after the third loan. By the end of March, 4,890 loans and 164 purchases had been recorded. It was clear that the service was popular, and feedback was almost universally positive, including comments such as: "This is cool. Much neater than Google books..."

As demand rose, it was felt necessary to further modify demand. The number of loans per student per day was reduced from three to one, with auto purchase on the fifth rather than the third loan. The cost to the library was also made visible – not with the intention of deterring users from requesting loans, but to make them more aware of what was being paid on their behalf. Whilst these measures slowed down spending, by May it was clear that further action was required so it was decided to mediate all requests. By the end of July, £86,000 had been spent on 7,300 loans and 254 purchases. Users had browsed over 30,000 titles and every academic school had made use of the service. Analysis of the user base revealed humanities & social sciences to be the dominant user group, and within this, final year dissertation students were particularly heavy consumers. The School of English, itself in the top ten user group, reported a reduction in the number of student complaints about access to books, and lecturers felt that EBL had played a part in this.

The results of the initial experiment gave the library confidence to continue the service in 2010/11. Once a full academic year's worth of usage data is available, there will be much to reflect on. The books selected by users will be analyzed to see if they appear on reading lists. Items selected for purchase by users will be monitored to see if they attract future use and will be compared with the usage of material selected by academic or library staff. The results of the next round of the NSS are awaited with interest and it is hoped that there will be evidence that the investment in patron-led e-book acquisition has met at least some of the student demand for more books. The challenge moving forward in a much more fiscally constrained environment will be to retain and fully embed an element of patron-led acquisition within the traditional acquisition model.

The morning session concluded with *Selena Killick's* presentation, **Evaluating the Big Deal: what metrics matter?**

In April 2010, Cranfield University Libraries embarked upon a review of their electronic journal packages following an indication of budget cuts. Previous budget reductions had led to cuts in book budgets and cancellations of journal titles from smaller publishers. It was anticipated that the potential shortfall would require a cancellation of one of the four large journal packages. Initial discussions focused on the CPD. However, it became clear that further detailed analysis was needed to understand the impact of any cancellations.

In line with a wider Library strategy to introduce key performance indicators (KPIs) to the overall service, KPIs for the electronic resources were sought. Eventually, the work conducted by Evidence Base in their Analyzing Publisher Deal project¹ and the KPIs implemented at Newcastle University Library² were adopted as the basis of performance indicators applied at Cranfield³. Using COUNTER JR1 reports and a three-year financial report from the Library's subscription agent, formulas were matched on ISSNs in an Excel spreadsheet to link title prices with usage and holdings. Three main output sheets were designed to meet the different management information needs; print subscriptions, top 20 and KPIs.

The print subscriptions report presented the three-year trends in title costs, downloads and CPD for the titles within the packages that were also taken in print format. Print usage data was not available so this was excluded from the report. Anecdotally, usage of the electronic journals is higher than the print journals; this known weakness in the data was seen as acceptable. Titles with low downloads and high CPD became a priority for cancellation or substitution.

The top 20 report showed the highest used titles within each package. Three calendar years of COUNTER JR1 reports were reviewed to show the titles consistently being used heavily, linked with the three-year pricing report to show if highest used titles were purchased in print format. Titles consistently in the top 20 that were not taken in print format became a priority for substitution or subscription. Cost data from the package provider was added to show the replacement costs for the highly used titles, along with the subject areas the title supported to give an indication of the

research areas that would be impacted by any cancellation.

The KPI report provided an overview of the basic and value metrics enabling the packages to be compared against one another. All of the metrics helped inform the review process, however the one metric that held more weight than any other was the subscription end date. The initial focus of cancellation discussions was on what *could* be cancelled rather than what *should* be cancelled. Out of the four packages under review, one had a cancellation date within the forthcoming financial year. Although cuts were made to the resources, the Library successfully lobbied to renew the package deemed at risk for one more year.

Looking ahead, the Library now intends to develop a systematic review process for all resources to demonstrate smart procurement. Although CPD is a useful indicator it will not be the sole metric used to assess resources. All metrics are important, but statistics only provide a two-dimensional view and should never be used in isolation. Other factors to be taken into consideration include qualitative local knowledge on the value and impact of titles within the collection.

The seminar presentations concluded with *Sarah Thompson's* **Counting the cost: the role usage statistics can play in a library subscription review**, a case study of how a subscription review was carried out at the University of York in 2010, with the aim of making £150K savings for 2011. Quantitative data was used alongside qualitative feedback to reach final cancellation decisions.

The academic community at York considers the journal collection to be integral to its research and teaching (teaching is largely research-led), and has high expectations that Library collections should be both 'broad and deep'. For a relatively small but steadily growing university (c.12,000 FTE in 2009/10), it has been increasingly difficult to meet these expectations in recent years.

There is a continual demand for more journal subscriptions, but growing student numbers and heightened student expectations have also increased the need for more books, so cutting the book budget was not a viable alternative to cutting the subscriptions budget. Most of the one-off purchases budget is spent buying books (both print and e) on reading lists, so cuts in this area have a real and immediate impact on student provision.

The Library has also experienced the same external pressures as other academic libraries in the UK in the last few years: high annual inflation on journals; adverse exchange rates; the global financial situation; UK HE funding cuts. The amount available to spend on subscriptions no longer matched commitments; cuts were needed, particularly if there was to be any hope of taking out new subscriptions to meet new needs.

Early in 2010, the Library began compiling usage data. This became a major task for the E-Resources Co-ordinator as the Library had not previously collected and analyzed usage data in a systematic way. Excel, Access and UStat (the statistics storage and analysis tool from Ex Libris) were deployed; usage data for calendar years 2008 and 2009 was consolidated, use of separately purchased (non-subscription) backfile titles was stripped out and details of subscribed titles and costs added in.

Following involved discussions about the criteria to be used to identify titles to cancel, and concerns that the library could not guarantee the usage data to be 100% accurate and that print usage figures were not available, staff were unwilling to take decisions on usage data alone. Rigidly applying other 'arbitrary' criteria (e.g. cost per use exceeds £5 and the subscription cost is greater than £50 and the total number of downloads is less than 10) was felt to be counter-productive.

A spreadsheet, including 3,519 subscription titles (excluding titles tied up in multi-year deals which could not be cancelled for 2011) and cost-per-use data was sent out to academic departments for feedback. To make it easier, titles were also coded as red, amber or green. Green titles were those with 200 or more downloads, which the Library expected to retain; amber titles had fewer than 200 downloads but Academic Liaison Librarians had suggested they be retained; both amber and red titles would be cancelled unless strong academic cases for their retention were received by departments. Academics were asked to rank titles in their subject areas on a scale of A–D, with A being essential, B important, C desirable and D no longer needed.

A special meeting of Library Committee followed, which almost all departments attended, in which academics emphasized the strategic importance of a strong journal collection to the University's research activity and postgraduate recruitment. Following this meeting, the Library Director agreed to reduce the cut to the 2011

subscription budget from £150K to £100K, with the shortfall to come from the book fund – but for one year only.

Despite the resulting cancellations, Sarah noted that the Library remains in a position many other libraries will identify with; recognizing the value of journals to academic activity and the success of the University, but without the budget to enable it to buy what is wanted. Strategic discussions continue with the University about Library funding. The Library expects to have to reduce its subscriptions portfolio further, unless it can get better deals and/or pricing from publishers. Consequently, the major review scheduled for 2011 will be even more significant, because it will include all the 'big deals', including those which were not considered in 2010 because the Library was committed to them until the end of 2011.

Sarah concluded by highlighting the fact that a large saving had been achieved, but at a cost! A cost in terms of the huge amount of staff time that had been involved – Project Team, Academic Liaison Librarians, Academics and serials staff – and she thanked all involved.

Throughout that day, it was clear that usage statistics can be an extremely useful tool in informing the decision-making process, but that they must be used carefully and with due consideration of the uncertainties and inconsistencies inherent in them. And it was equally clear that no one yet has a simple, foolproof answer to how to use them effectively.

If you would like to read what one attendee thought of the seminar, see Sheree Dewey's comments in *Serials-eNews*, Issue 234, 26 November 2010.

References

1. Conyers, A, What do publisher usage statistics tell us? The Analysing Publisher Deal project from Evidence Base, *SCONUL Focus*, 2007, 40, pp.72–76.
2. Newcastle University Library, 2010. Lib-Stats: Key Performance Indicators 2007 – Journals [online]: Available at: <http://www.lib-stats.org.uk/> (accessed 03 December 2010)
3. Killick, S, *Evaluating the Big Deal: What Metrics Matter?* Presented at UKSG Usage Statistics for Decision Making. London, 18 November 2010. Available at: <http://www.uksg.org/sites/uksg.org/files/PresentationKillick.pdf> (accessed 03 December 2010).