

# A librarian's view of usage metrics: through a glass darkly?

*This article is an extended version of a paper presented by Jill Taylor-Roe at the 28th UKSG Conference, Edinburgh, April 2005*

A research project at Newcastle University Library has been benchmarking electronic resources using COUNTER-compliant usage data. As the virtual academic library becomes more widely established, new tools will be needed to determine the cost benefit of library resources and services. In this context, traditional quantitative data will no longer suffice. E-metrics have the potential to provide managers with more robust, empirical evidence to inform strategic planning and decision making.



**JILL TAYLOR-ROE**  
Head of Liaison &  
Academic Services  
Newcastle University



**CLIFF SPENCER**  
Engineering Liaison  
Librarian  
Newcastle University

## Context and background

As library managers, we spend a considerable amount of our time collecting, analysing and acting upon usage data. Why do we devote so much effort to this activity? Usage data is a tool that can help us manage our libraries effectively, and it can do this in several ways.

### ■ It can highlight trends

For example, sustained, significant increases in reservations could trigger a review of loan periods and levels of copy provision, a consideration of e-access over print, or could trigger further dialogue with academic staff (did they remember to send us the reading list?). E-journal usage data not only confirms that just a relatively small proportion of our journal titles have significant use but, especially within the big deal packages, enables us to monitor and measure demand for journals which are beginning to emerge as important titles, and which may highlight new or developing research interests.

### ■ It can inform our strategic decision making

For example, we might say that when self-issue statistics account for 60% of all issues, we will redesign the front counter and deploy some of the front line library staff in other capacities. With regard to e-metrics, we might decide to

utilize turnaways to inform collection development decisions regarding the purchase of additional or alternative resources, either as subject clusters or individual title subscriptions.

Usage data can also help us to assess whether our library services are meeting user needs. One of the absolute certainties in any user consultation exercise we carry out at Newcastle is that undergraduates will ask for more textbooks and postgraduates and academic staff will ask for more journals. If we acquire new resources in response to these expressed needs, we automatically look at the subsequent usage data to see if these new acquisitions are being utilized and we elicit feedback from the initial requesters to see if their needs have now been met.

Of course there is also a more pragmatic, political value to usage data. We use it in our annual reports, our business plans and as a fundamental component of our regular dialogue with the University to demonstrate that we are providing value for money and to reassure our paymasters that the resources entrusted to us are being effectively deployed in support of teaching, learning and research. We also use it to back up requests for further investment, or to support bids for external

funding. Whilst a significant amount of the usage data our library collects is simply measuring the level or volume of activity, we also acquire and analyse a wide range of qualitative data. For example, we regularly solicit feedback on what our users think about our services. What are they happy with? What are they dissatisfied with? What would they most like us to do differently? Every three to four years, Newcastle University Library commissions a major survey of user needs and views which is carried out by a company called Priority Research. The information we gain from the survey is used to inform our subsequent planning strategy. Some of the most practically useful data is that which we gain from responses to questions where users have to select between paired examples of comments they have made in the preliminary focus groups. In this way, the Library ends up with a prioritized user 'wish-list' which is much more practical to work with than a long list of 'wants'.

Usage data which simply relates to volume or level of activity is invariably one-dimensional and seldom tells the whole story. Nevertheless, much of the data traditionally collected by academic libraries falls into this category. For example, here are four very typical pieces of library usage data which have been collected over several years for the Society of College, National and University Libraries (SCONUL)<sup>1</sup> annual statistical return.

The issues referred to in Table 1 are essentially book issues. Current parts of our print journals are reference only, and although precurrent bound volumes of journals are available for loan, they account for a very small proportion of total loans. Our issue statistics show a peak of just over 760,000 in 1999/2000 but they are now back to just below 1994/1995 levels.

ILL (inter-library loan) requests were almost 36,000 in 1994/1995, but by 2003/2004 this had plummeted to just over 13,000. This represents a decline of 63.2%. Similarly, photocopying is down from a high of almost 4.5M in 1999/2000 to 1.74M in 2003/2004 – a decline of 61.2%. This fall in demand has had a significant impact as, traditionally,

photocopying income has been used to pay for photocopying staff and equipment, as well as small-scale projects which the Library could not readily accommodate within its Block Grant expenditure.

Interestingly, although there has been a 20% decline in library visits over the ten-year period, there has not been the steep decline that some of the other usage measures have witnessed. Perhaps this is because users still view the Library as an important social space, or that we have the largest computing cluster on campus, with the longest opening hours. Certainly, the customers are still coming through the doors, but what are they doing when they get here?

Clearly, if we were relying solely on these traditional usage metrics to demonstrate the continuing effectiveness of the Library, our paymasters could perhaps be forgiven for wondering whether they had grounds to reduce the Library budget. Fortunately, we have new and additional metrics to demonstrate that library activity is changing, and indeed expanding, rather than declining. Although we have never sought to prove it, it is generally believed that the huge decline in ILL requests and photocopying is inextricably linked to the emergence of e-journals. The Pilot Site Licence Initiative (PSLI)<sup>2</sup> brought the first three publisher collections into academic libraries in 1996/1997 and at Newcastle we now provide access to over 9,000 e-journals, many of which have been secured via the National Site Licence Initiative (NESLI) and its successor, NESLi2.<sup>3</sup>

### Expenditure on e-resources

As this paper is going to focus on the analysis of usage data pertaining to e-resources it would be helpful to contextualize this. At Newcastle we are currently spending around £1.3 million (55%) of our resources budget on electronic resources. With this we are providing access to over 9,000 e-journals, 190 databases and around 500 e-books.

Academic Year	94/95	99/00	03/04
Issues	574,618	760,128	505,044
ILL requests	35,771	23,335	13,180
Photocopies	4,431,397	4,498,000	1,745,648
Visits to the Library	1,243,970	1,175,249	1,012,670

Table 1. Newcastle University Library standard usage data

The expenditure and the available e-content are increasing significantly each year. Whilst a small proportion of this expenditure has come from additional project funds, the bulk of it is Block Grant expenditure, which means that the more we spend on e-resources, the less we have available to spend on print or indeed on any of the other information formats which our users require. (For example, we are still asked to buy microforms for some of our humanities researchers.) Demand for new resources in all formats is insatiable, so it is of paramount importance that we know, and are able to demonstrate, that the resources we choose to buy are earning their keep. This is particularly important in relation to e-journals where the dominant mode of purchase is now the package – either the big deal, or the subject cluster. However, when academics ask for new journals, they invariably ask for individual titles, not packages. Traditional library collection development policies were founded on a 'cancel one to buy one' approach – a model which does not function well (and often not at all) in the package deal environment. Academics are often frustrated when they see journals in our e-collections which they do not require or value, and which they suggest we cancel to fund the additional titles they need, only to be told that the unwanted titles are part of a package and cannot be exchanged for titles from another publisher.

### Benchmarking at Newcastle

Thus from a library perspective, we need to be confident that the entire journal package is performing well, and we are just as interested in the extent of the tail of 'low to no' use as we are the top performing titles. To this end, we are using local benchmarking techniques to determine the relative performance of our e-journals, and have developed a range of associated key performance indicators (KPIs) that we now apply to all our e-journal packages. Table 2 shows the range of KPIs Newcastle is currently using for 2005. The nature of the benchmarking process is such that KPIs are constantly reviewed, and modified where necessary. The KPIs are populated, wherever possible, with data that is COUNTER compliant.

Monthly usage statistics are collected from each service provider and downloaded into a detailed spreadsheet based on the above-listed metrics. The spreadsheet is used to produce a range of graphs to

#### Metrics

Total number of titles in package  
 Total full-text downloads  
 Max. downloads (month)  
 Min. downloads (month)  
 Mean downloads (year)  
 PDF full-text downloads  
 HTML full-text downloads  
 PDF/HTML ratio  
 Downloads per FTE user  
 Total turnaways (titles not included in licence)  
 High use ( $\geq 100$  hits)  
 Medium use (11–99 hits)  
 Low use (0–10 hits)  
 Top 20 best performing titles as a % of total hits

#### Costs

Total package costs (= e+p where appropriate)  
 Cost as a % of total serials budget  
 Mean cost per title  
 Mean cost per full-text download  
 Cost per FTE user

#### Additional Metrics

Total pay-per-view (PPV)  
 Total images  
 Top 20 best performing titles – subscribed v. non-subscribed  
 Top 20 best performing v. impact factor  
 Subject distribution of titles by usage  
 Top 10 most expensive titles v. usage

#### SFX clicks per title

Table 2. Key performance indicators (KPIs) 2005

help us assess the performance of individual journal packages. Data is derived from COUNTER-compliant sources and merged into one discrete set.

We are beginning to compare relevant publisher packages, for example STM, social sciences, humanities, so that we can see how packages with broadly similar content are performing.

The graph represented in Figure 1 illustrates how full-text downloads have increased over a three-year period. Usage is generally higher in term-time, with noticeable peaks at the start of the academic year in October/November, and also in February/March, after the January exams are over. We are also finding that whilst, overall, usage is increasing year on year, the usage patterns for more established packages become more regular, and the annual increases are less dramatic, as shown in Figure 2.

We plan to carry out more research to determine why the usage patterns become more settled.

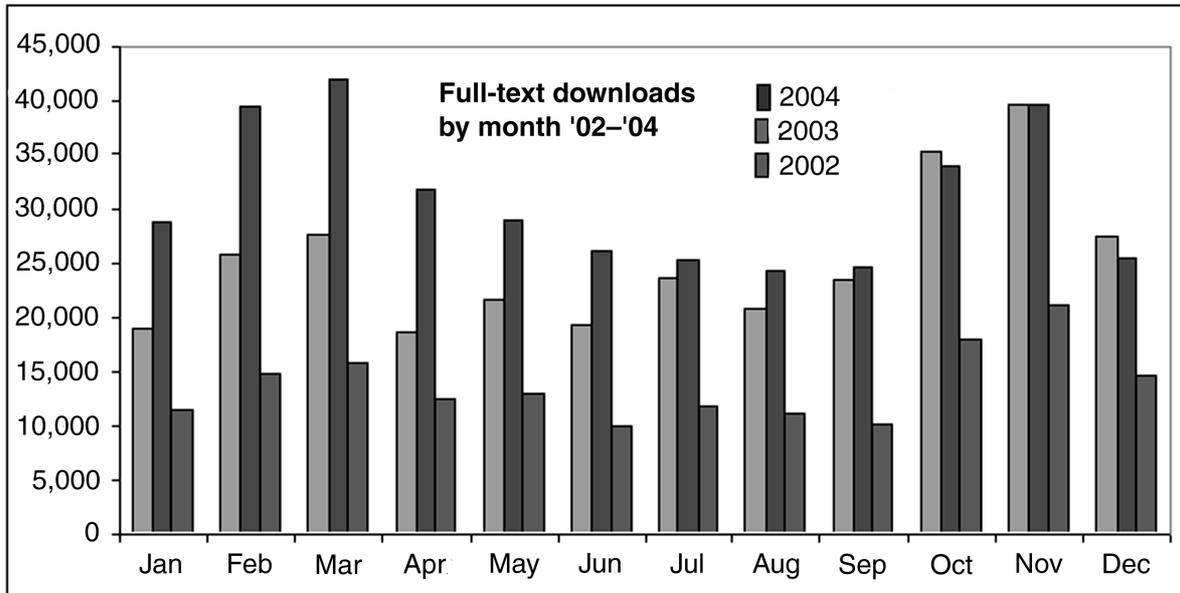


Figure 1. Monthly full-text downloads for a single STM package

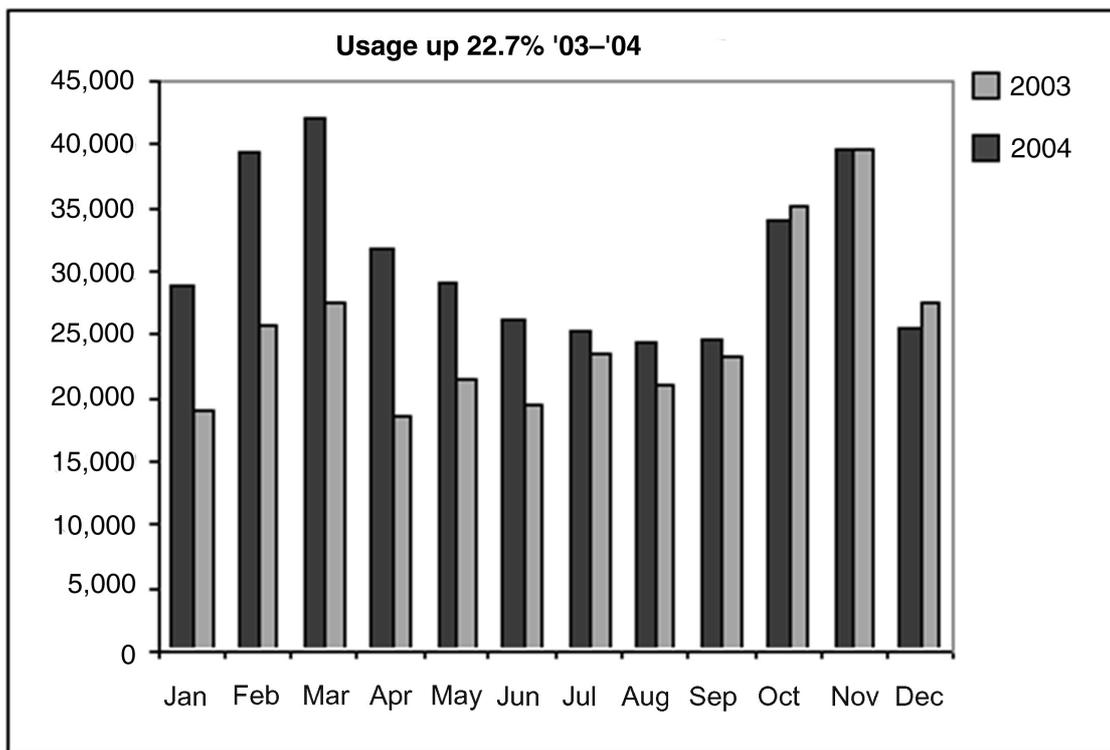


Figure 2. Annual downloads for a well established package

However, as shown in Figure 3, with newer packages the increases can be much more striking, particularly if we have carried out some additional promotion, or have bought the package as a direct result of expressed demand. For example, New-

castle has recently taken out a subscription to *Nature* and the associated *Nature Research* titles and we will be monitoring the usage very carefully to see how it compares with other similar, newly acquired, collections.

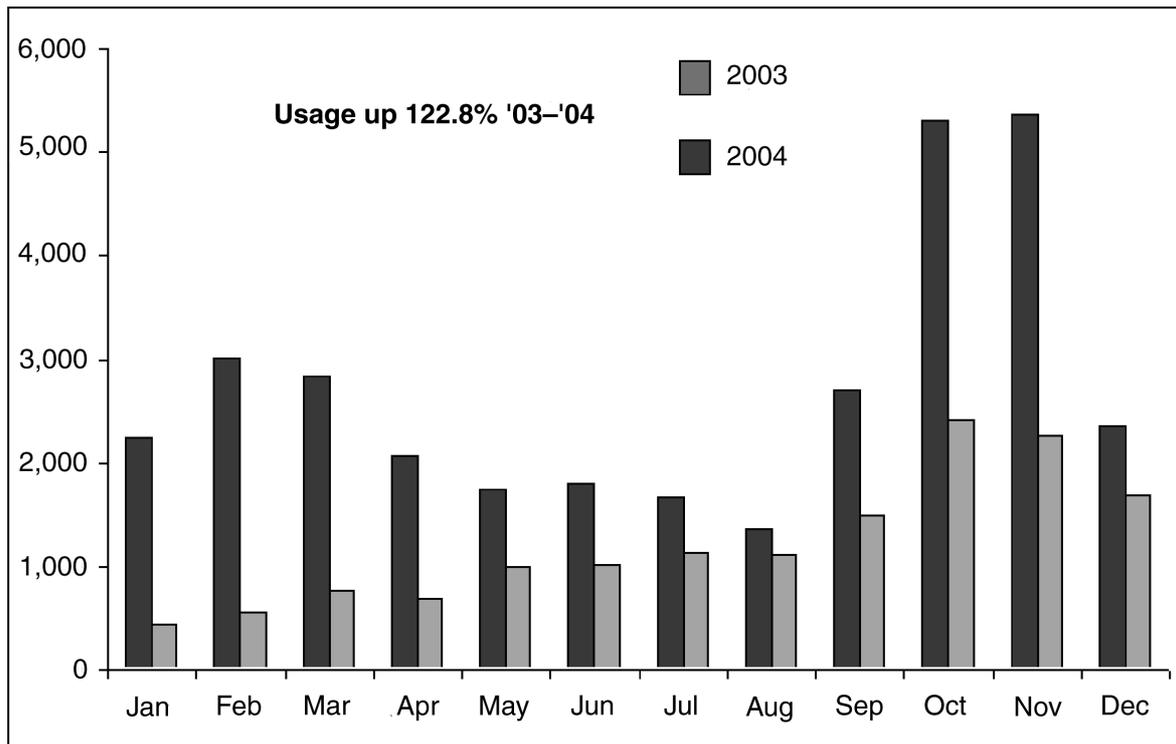


Figure 3. Percentage change in annual usage for a newly acquired package

### Additional research findings

What else are we learning from this research?

- STM titles dominate 'high' usage rankings.
- Our two best performing STM packages have increased usage year on year for the last three years, and have also seen significant reductions in the percentage of titles reporting no usage (now down to less than 10% of the total content).
- Subscribed titles are generally more heavily used than unsubscribed.
- Typically, a relatively small number of journals in a large package account for a significant percentage of total downloads.
- Sustained low usage of individual titles or subject clusters can be used as a justification for cancelling content, where licence terms permit. Similarly, sustained evidence of turnaways can be utilized to justify buying additional or alternative subject clusters.

We are starting to add impact factors to the STM collection usage data to see if high impact factor titles correlate with high use. We also intend to explore whether we can assign the appropriate RAE subject codes against titles as we feel it would be interesting to see how the different subject areas perform within a large package.

We are looking closely at the cost per full-text download, which is generally decreasing and for our major STM packages is currently between £1.77–£2.00. This compares very favourably with traditional ILL costs which Birch and Young<sup>4</sup> in their 2001 study estimated at £13.30. Ideally, further research is needed to determine the current costs (perceived and hidden) of an inter-library loan, and we felt that further attention should also be paid to the true comparative costs of a full-text download.

So why are we devoting so much attention to this data? We want to develop more cost-effective ways of determining and evaluating our e-journal purchasing activity to ensure that we are making the best use of available funds. We think there is a comparatively small core of big deal packages that are truly essential for Newcastle. Around this inner core will be a band of desirable subject clusters; around this layer are the titles we only need to acquire as individual subscriptions; and finally, the outer core will be the occasional articles which we may decide to acquire on a pay-as-you-go basis. Given that our funds are finite, but demand for them is infinite, we believe that e-metrics can help us to assess when it is most cost effective to pursue a particular purchasing model.

### Future work

As we build up several years' worth of comparative usage data, it should become much clearer when a big deal is not working. For example, if after three to five years, the tail of low to no use has remained comparatively constant, or represents a significant percentage of the total content, or if the cost per full-text download is much higher than that of comparable packages, then we would start looking at the analysis of usage by subject to see if a subject-based approach might be more appropriate.

We are beginning to use the data we gather to replace or enhance some of the traditional library metrics. Thus the decline in ILL and photocopying can be offset against the significant growth in full-text downloads per annum. More recently, we are finding that e-metrics are a very helpful component in dialogue with academic colleagues in the development of new modules. In common with many universities, new modules or degree programmes at Newcastle have to go through a rigorous assessment to prove that there is a sound business case to justify the investment and start-up costs. Part of this assessment involves the library costing out the required information resource provision.

For example, we have an MSc in Oncology and Palliative Care which was set up in 2002 as a three-year pilot. It is a distance learning course for clinicians and we have spent around £30,000 over the three years securing subscriptions to a suite of e-textbooks together with a few additional journals which were specifically requested for the course. The course started off with 20 students, and now has over 70. As part of our evaluation, we have been able to feed back to the course co-ordinators not just how much we have invested in supporting the module, but how well the resources have been used. Obviously, there are caveats with the reports; we already had strong collections in this subject area, so the extra resources we have bought are not the only ones the students are using. Furthermore, the resources we bought for this course could also be used by other bona fide users within Newcastle University. Nevertheless, it is still extremely important to know, and be able to demonstrate, that we are getting value for money in relation to the additional investment that was made to support this course.

We are now preparing to look critically at our database collection and apply some of the more

rigorous e-metrics we have developed for e-journal usage to them. As our full text journal collections have grown, there is certainly a perception amongst the liaison librarians that some of our users are ignoring the subject databases and going straight to the e-journal collections. We suspect that, as with the big deal journal collections, there is a small core of databases that are truly essential for us, but in other more specialist subject areas, it might be more cost effective to pay for online searches on demand. We need to look critically at the usage data and costs of these databases over a period of years to see if there is evidence to support these views.

We will also assess whether the implementation of MetaLib, which we are planning to launch this summer, has a positive impact on database usage. We already have SFX up and running, but it is really too early to tell to what extent it is affecting access to e-journals. And as our e-book collection is finally beginning to expand, we are going to start developing some comparable e-metrics for them.

We will certainly do more work with turnaways, although it would be helpful to have more information from publishers on exactly what these mean in the context of their usage data. For example, COUNTER defines a turnaway as 'an unsuccessful log in to an electronic service due to exceeding the simultaneous user limit allowed by the licence.'<sup>5</sup> But it could also mean that there was system downtime or that the user was trying to access unsubscribed content. We need to know exactly which definition pertains in order to determine the most appropriate response.

We have just gone live with the pilot phase of our e-print repository and as it becomes well-populated, we will be looking closely at what is being used and how this relates to our subscribed content. This is of course but one small strand of the much wider open access movement and it is too early to assess the implications of that. Then there is Google Scholar – we really don't know what its impact will be on our usage data, but we know we need to be mindful of it, because of its inevitable mass appeal.

### Summary and conclusions

Based on our experience to date, e-metrics, although extremely time-consuming to collate, are already demonstrating their value. They provide us with

much more detail than our more traditional, print-based metrics. Once you have the data downloaded into a format that can be manipulated, it is possible to compare the performance of individual titles or subject areas; or to assess usage trends over particular months, semesters, or years. Of course it means more work, and that has to be appreciated and planned for, but e-metrics raise our expectations in ways which traditional, print-based usage data never did, and never could.

E-metrics can be useful in assessing the cost of supporting particular programmes or modules and in determining whether the usage made represents good value for money.

Let us not forget, usage metrics are still a work in progress. It is not so very long ago that where current print journals were reference only, library usage was assessed by clearing them off tables or photocopying trolleys and logging usage against a printout of journal titles. In the past, libraries also experimented with consultation records which were periodically attached to current issues of titles thought to be 'at risk' and which academics were asked to sign every time they used the journal. We have come a very long way already since those days, but of course we still want, and need, to go much further.

### Problems and challenges

It is very time consuming to prepare e-journal usage reports. The work is particularly challenging when you have to pull in data from aggregators as well as from the publisher's web site, and the lists of titles do not match. A considerable amount of time is then wasted trying to find out what is missing from one list, and why. The work of COUNTER has been invaluable in improving the overall quality of the data and ensuring greater standardization. We always notice a significant improvement in the quality of reporting when a publisher becomes COUNTER compliant.

It still seems extraordinarily difficult to determine a list of print subscriptions which the library and publisher can agree on as the basis for a package deal. Sometimes we go through several iterations before we reach agreement and without the help of our subscription agent this would be an even lengthier process. We must find ways of making this easier for everyone concerned.

The level of IT skills required for library staff to effectively collate and exploit usage data is also a concern. We are very fortunate at Newcastle that we have someone who can undertake this work for us, but dialogue with other libraries suggests this is not always the case. There is a limit to what can be done locally and I think our best opportunities for further progress are by means of national or possibly regional initiatives.

### Where next?

Certainly there would be huge benefits in having some centralized reporting for statistics – perhaps through COUNTER, or NESLI. It would also be extremely helpful if we could analyse usage by off-campus as well as campus use. This would not only be useful to those of us with distance learners, but could also help to demonstrate how the virtual library is working, and could be seen as an important performance indicator when visits to the physical library are declining.

We would also like to have usage data which showed which years of a journal title were being used. This would not only help in terms of assessing the merits of back-files but is of interest in assessing the value of current versus precurrent information in different subject disciplines. Many of these concerns are echoed in the recently published NESLi2 report by Conyers and Dalton.<sup>6</sup>

In conclusion, usage metrics have developed rapidly over the last three years, and will no doubt continue to develop at a similar pace. They are fundamentally important to us in our practical roles as library managers and also have an added political value within the organization. SCONUL has already recognized the growing importance attached to e-metrics and now includes questions on e-resources in its annual statistical return. There is much more we need to learn about e-metrics and we must ensure that expertise and good practice are shared. As we gain in experience and knowledge, we will seek and achieve greater refinements in the data we collect and analyse.

As has been demonstrated by the COUNTER initiative, a collaborative approach can bring significant benefit to all interested parties and this is likely to hold true for subsequent developments in e-metrics. Certainly, if we are to develop benchmarking to its logical conclusions, where we can

compare our usage trends with those of other libraries in order to assess the wider picture, we must ensure that the metrics we are using are of a comparable standard and format.

The sub-title of this article, 'through a glass darkly'<sup>7</sup>, suggested that as far as e-metrics were concerned, we still had some way to go before we had achieved total clarity. That still seems to us to be a fair assessment. However, as the novelist George Eliot memorably observed, "Beginnings are always troublesome"<sup>8</sup>. She certainly felt it was worth persevering – and so do we.

## References

1. Society of College, National and University Libraries (SCONUL)  
<http://www.sconul.ac.uk/>
2. Report on Phase I of the Evaluation of the UK Pilot Site Licence Initiative  
[http://www.hefce.ac.uk/pubs/hefce/1997/m3\\_97.htmress](http://www.hefce.ac.uk/pubs/hefce/1997/m3_97.htmress)
3. National Site Licence Initiative  
<http://www.nesli2.ac.uk/>
4. Birch, K. and Young, I.A.  
Unmediated document delivery at Leeds University: from project to operational system. *Interlending and Document Supply*, 2001, (1) 4–10.
5. COUNTER Project COUNTER (Counting Online Usage of Networked Resources)  
[http://www.projectcounter.org/r2/Appendix\\_A\\_Glossary.pdf](http://www.projectcounter.org/r2/Appendix_A_Glossary.pdf)
6. Conyers, A. and Dalton, P.; NESLi2. Analysis of Usage Statistics: Summary Report.  
[http://www.jisc.ac.uk/uploaded\\_documents/nesli2\\_usstudy.pdf](http://www.jisc.ac.uk/uploaded_documents/nesli2_usstudy.pdf)
7. *The Holy Bible*, 1 Corinthians, ch. 13, v.1
8. Eliot, G., letter to Sara Hennell, 15 August 1859 in G.S. Haight (ed) *The George Eliot Letters* (3), 1954, New York, Yale University Press.

Article © Jill Taylor-Roe and Cliff Spencer

---

■ **Jill Taylor-Roe, Head of Liaison & Academic Services**  
Tel: +44 (0)191 222 7652  
E-mail: [Jill.Taylor-Roe@ncl.ac.uk](mailto:Jill.Taylor-Roe@ncl.ac.uk)

**Cliff Spencer, Engineering Liaison Librarian**  
Tel: +44 (0)191 222 3425  
E-mail: [C.Spencer@ncl.ac.uk](mailto:C.Spencer@ncl.ac.uk)

**Robinson Library, Newcastle University**  
**Back Jesmond Road West**  
**Newcastle upon Tyne NE2 4HQ, UK**

---

*Article received 9 May 2005; revised 9 June 2005; accepted 13 June 2005; published online July 2005.*

To view the original copy of this article, published in *Serials*, the journal of the UKSG, click here:

<http://serials.uksg.org/openurl.asp?genre=article&issn=0953-0460&volume=18&issue=2&spage=124>

For a link to the full table of contents for this article, please click here:

<http://serials.uksg.org/openurl.asp?genre=issue&issn=0953-0460&volume=18&issue=2>