**World Library and Information Congress:**
**70th IFLA General Conference and Council**

**22-27 August 2004**
**Buenos Aires, Argentina**

***Programme: http://www.ifla.org/IV/ifla70/prog04.htm***

**Code Number:** 098-E
**Meeting:** 125. Statistics and Evaluation with University Libraries
**Simultaneous Interpretation:** Yes

# Managing the Electronic Collection with Cost per Use Data

**Brinley Franklin**
**University of Connecticut Libraries**

## Abstract:

*This paper reviews some of the early efforts to develop cost per use data for electronic collections and discusses some of the ways libraries, consortia, and publishers currently use unit cost information to make management decisions. Emerging trends in the standardization of electronic usage statistics and concurrent utilization of cost per use data to manage electronic collections hold tremendous potential for libraries and library consortia to increasingly employ reliable cost and use data to support collection development and management decisions.*

The pricing of networked electronic resources is still in its early stages of development. Libraries often purchase electronic journals differently than the traditional method of an annual subscription cost for a print journal. Despite promising recent initiatives like COUNTER (Counting Online Usage of Networked Electronic Resources)[1], there has historically been too much variation in vendor usage reports to accurately compare cost per use data for electronic resources across vendors and publishers. Some libraries compute unit costs solely on the basis of subscription costs. Fewer libraries take a more comprehensive approach and compute the total cost to offer electronic collections to users.

Despite variations in pricing models for acquiring electronic journals and the lack of consistency to-date in usage reports from vendors, the ability to measure usage of networked electronic

resources holds great promise for librarians as a tool for measuring not just the usage of electronic resources, but also the unit cost of each database search or article view or download. This paper reviews some of the early attempts to develop cost per use data for electronic collections and discusses some of the ways libraries, consortia, and publishers can use unit cost information to make management decisions.

**Unit Costs of Electronic Versus Print Journals**

To manage the transition from print to electronic journal collections successfully, library managers could examine how the unit cost of electronic journal usage compares to print journal use. An American library, Drexel University Library in Philadelphia, received funding from the U.S. Institute of Museum and Library Services, beginning in 2000, to study the economic effects of its library's shift from print to electronic journals on staff and other costs.[2] The Medical Branch Library of the University Library in Muenster, Germany studied the correlation of usage in the same set of print and electronic journals and, in part, sought to determine the cost efficiencies of print and electronic journals.[3]

By 2002, Drexel University had migrated to an almost exclusively electronic journal collection. In 1998, Drexel subscribed to 1,710 print journals and 200 electronic journals. In 2002, it subscribed to 8,600 electronic journals and 370 print journals. The Drexel University study is especially important because it considered all costs associated with journal subscriptions (i.e., the subscription price as well as operational costs). Operational costs included: space costs required to house print collections; systems costs (e.g., servers, workstations and software, including maintenance costs); supplies and services (e.g., binding; security strips, printing); and staff costs

Operational costs to service electronic journal collections include compensation for staff whose skills include licensing, managing the library's offerings in journal packages whose content may quickly change, implementing and maintaining journal linking software, and other staff costs not required in a print journal environment. Annual operational costs for print and electronic journals at Drexel University were reported as follows:

| | *Electronic Journals* | *Current Journals* | *Bound Journals* |
|---|---|---|---|
| Space | $5,000 | $40,000 | $205,000 |
| Systems | $10,000 | $2,500 | $2,400 |
| Supplies and Services | ($2,000) | $600 | $8,000 |
| Staff | $125,000 | $46,000 | $42,000 |
| *Total Operational Costs* | *$138,000* | *$90,000* | *$258,000* |

When determining e-journal subscription costs, Drexel found that pricing models and content offered varied so significantly among different types of electronic journal that four electronic journal subscription categories were established: individual subscriptions; publishers' packages; aggregrator journals; and full-text database journals. Costs for electronic journals were based upon either individual subscription price or the average cost per title in a publisher's package, an aggregated journal package, or a journal offered in a full-text database. In Drexel's pricing model, only half the cost of a database was included since databases serve a larger function than

just providing the full-text of journals. Drexel reported the following subscription costs for the four different journal categories in 2002:

| | *Titles* | *Total Cost* | *Cost per Title* |
|---|---|---|---|
| *Print Journals* | 370 | $38,000 | $100 |
| *Electronic Journals* | | | |
| Individual Subscriptions | 266 | $115,000 | $432 |
| Publisher's Packages | 2,500 | $334,000 | $134 |
| Aggregator Journals | 480 | $29,000 | $60 |
| Full-text Databases(non-unique) | 10,200 | $59,000 | $6 |
| *Unique Electronic Journals* | *8,600* | *$537,000* | *$62* |

To determine unit cost information for print journal titles, Drexel maintained title-by-title re-shelving counts for bound and unbound journals for four years. Electronic journal use was defined as an article view, accessing an html file, or downloading a PDF file. Montgomery and King noted that print and electronic use data are not directly comparable for several reasons. Most notably, print journal use based on re-shelving assumes only one article was referenced before it was re-shelved while each use of an article within the same electronic journal issue by even the same user is recorded as a distinct use.

To further compound matters, neither re-shelving counts nor electronic use data received from publishers reflect truly accurate counts of journal usage. Davis and others have written why electronic use cannot be compared to print use, arguing that print journal use may be underestimated by as much as 80% and that electronic journal use data reported by publishers should be treated with skepticism.[4]

Nevertheless, Montgomery and King presented the following summary of cost per use by journal type for those titles where usage data was available:

| | *Recorded Use* | *Subscription Cost per Use* | *Operational Cost per Use* | *Total Cost per Use* |
|---|---|---|---|---|
| *Print Journals* | | | | |
| Current Journals | 15,000 | $2.50 | $6.00 | $8.50 |
| Bound Journals | 9,000 | n.a. | $30.00 | $30.00 |
| *Total Print Journals* | *24,000* | *$2.50* | *$15.00* | *$17.50* |
| *Electronic Journals* | | | | |
| Individual Subscriptions | 23,000 | $3.20 | $0.45 | $3.65 |
| Publisher's Packages | 134,000 | $2.25 | $0.45 | $2.70 |
| Aggregator Journals | 20,000 | $1.35 | $0.45 | $1.80 |
| Full-text Database Journals | 158,000 | $0.40 | $0.45 | $0.85 |
| *Total Electronic Journals* | *335,000* | *$1.40* | $0.45 | *$1.85* |

Montgomery and King acknowledged that usage data was problematic. They pointed out, however, that these large differences in unit costs were meaningful and reached the following conclusions as a result of the Drexel University study:

- Operational cost per use for print journals ($15) was much greater than for electronic journals ($0.45);
- The highest cost per use ($30) was associated with bound journals, given the cost to house them and their relatively low use;
- Full-text database journals were used heavily and were cost effective (at less than $1 per use); and
- Unit costs for publisher's packages and aggregator journals were more cost-effective than individual subscriptions, but further investigation of the use patterns within these packages is warranted.

A subsequent study at the Medical Branch Library of the University Library in Muenster, Germany also reported a significant difference in the unit cost of an electronic journal versus a print journal. One component of the study reported by Obst in 2003 was to differentiate the cost efficiencies of print and online journals.

Like Drexel, the ratio of print journals subscriptions to electronic subscriptions changed rapidly between 1998 and 2002. In 1998, the Medical Branch Library subscribed to approximately 700 journals online and 800 print journals. In 2002, the Library offered 1400 journals online and about 600 print journal subscriptions. In the Muenster study, print and online usage was compared for 270 journals for which both print and online versions were available to users.

The Muenster study only considered subscription costs. The definition of subscription costs has become problematic as vendors moved in recent years from print subscriptions to print plus optional electronic subscriptions to electronic plus optional print subscriptions. To account for subscription costs that included both print and electronic journals, the Muenster study included prices quoted by their publishers for an exclusively print or electronic subscription.

Print usage data in the Muenster study was collected quite differently than in the Drexel study. In the Muenster study, print usage data reflected total copy volume from which monograph copies (25%) were subtracted. The net number of journal copies was then divided by ten (based on a methodology adapted from Bauer)[5] to estimate the number of articles copied.

The Muenster study reflects significant differences in unit costs not just for print versus electronic subscriptions, but also between publishers:

| *Publisher* | *Print Subscription Cost (Euros)* | *Print Usage* | *Print Unit Cost (Euros)* | *Online Subscription Cost (Euros)* | *Online Usage* | *Online Unit Cost (Euros)* |
|---|---|---|---|---|---|---|
| Academic | 67,533 € | 3,350 | 20.16 € | 54,241 € | 3,593 | 15.10 € |
| Blackwell | 35,742 €. | 1,531 | 23.35 € | 30,380 € | 6,329 | 4.80 € |
| Elsevier | 60,143€ | 4,012 | 14.99 € | 54,139 € | 8,248 | 6.56 € |

| | | | | | | |
|---|---|---|---|---|---|---|
| HighWire | 8,984 € | 3,223 | 2.79 € | 8,086 € | 25,975 | 0.31 € |
| Springer | 85,335 € | 1,679 | 50.82 € | 76,801 € | 20,346 | 3.77 € |
| *Total* | *257,737 €* | *13,795* | *18.68 €* | *223,647€* | *64,491* | *3.47 €* |

The Muenster study demonstrated that, for titles included in the study, users not only accessed the online version significantly more than the print version, but also, on average, the unit cost of an online usage was 5.38 times (18.68 €/3.47 €) cheaper than the unit cost of print journal usage.

**Using Unit Cost Information for Management Decisions**

To make comparisons between electronic resources for the purposes of collection development or content delivery to a particular constituency, individual libraries currently calculate a relatively straightforward cost per use figure for their networked electronic resource offerings based solely on subscription costs. The temptation to perform and employ this simple calculation seems appealing because the electronic information environment yields a vast amount of easily manipulated usage information, particularly for electronic journals, that the print journal environment did not.

At the University of Virginia, electronic journals provided by twenty-six publishers and aggregators cost the University Library about $2.3 million in subscription costs in fiscal year 2003. When print and online costs were bundled, 75% of the cost was assigned to e-journals and 25% to print journals. There were 1.4 million recorded uses of the articles contained in these electronic journals, yielding an average cost per article used of roughly $1.64. There was a significant range of cost per article use among the twenty-six publishers and aggregators, however. Cost per article used varied from $.07 to $17.92; the median publisher/aggregator cost per article use was slightly more than $1.00.

Forty-nine of the major electronic reference databases to which the University of Virginia subscribes cost approximately $550,000 annually. These forty-nine databases were searched about 1.05 million times last year, yielding a $0.52 per search cost. The most heavily used database was searched more than 270,000 times, at a per search cost of $0.08. An infrequently searched database had the highest per search cost ($16.68).
A number of universities perform cost per use calculations in a similar way. Another American university, Western Carolina University, reported the employment of cost per use data as "a valuable decision-making tool during a periodical review cycle when the need arose for identifying a percentage of titles to cancel.[6]

At the University of Connecticut, users downloaded 180,521 articles from ScienceDirect in calendar year 2003. The list price of the journals we offer our users through ScienceDirect was approximately $2 million in 2003. Consequently, it cost the University, on average and using list prices, about $11 per article download. Our actual cost per article download was significantly less (about 60% of list price) because we did not pay Elsevier the list price, but rather a capped price for our former print subscription journals, plus a cross access fee for additional journals we did not subscribe to in print, plus an electronic access fee.

Our ScienceDirect cost per use varied from $0.57 for a journal with a list price of $635 that was used 1,121 times to $1,230 for a journal that was used once. Moreover, we found that journals we had not subscribed to in the print environment were used fairly frequently when they became available as part of the ScienceDirect package and that some journals we subscribed to in print were not heavily used when offered electronically. Two physics journals that we subscribed to in print, for example, with list prices that total about $24,000, were used forty-one times. Our second most expensive journal was used more than 2,000 times, for a cost per use of about $9. Our most heavily used journal in Science Direct was used 3,275 times in 2003, at a cost of about $1 per use.

Librarians at the University of Connecticut can also rather easily determine the cost per search for many of our most heavily used databases because we have subscription costs and usage data readily available. This database comparison does not address the different content of these networked electronic resources; it only compares cost, usage, and the resulting cost per search figure for two consecutive years. This simple subscription cost per search calculation results in the following management information:

| | *FY 2002 Searches* | *FY 2002 Cost* | *FY 2002 Cost per Search* | *FY 2003 Searches* | *FY 2003 Cost* | *FY 2003 Cost per Search* |
|---|---|---|---|---|---|---|
| WorldCat | 35,762 | $9,766 | $0.27 | 30,131 | $11,160 | $0.37 |
| JSTOR | 31,485 | $23,100 | $0.73 | 49,439 | $30,650 | $0.62 |
| FirstSearch | 14,736 | $9,380 | $0.64 | 13,700 | $6,375 | $0.47 |
| Web of Science | 72,040 | $135,270 | $1.88 | 66,420 | $144,039 | $2.17 |
| *Total* | *154,023* | *$177,516* | *$1.15* | *159,690* | *$192,224* | *$1.20* |

A simple example like this provides a library manager with various ways of looking at the rising cost of networked electronic resources. The cost of these four databases increased by 8.3% in one year. The number of searches performed on these databases increased 3.7% during the same year. The unit cost per search in these databases increased on average by 4.3% from 2002 to 2003. The expanded JSTOR database, whose price increased by the highest percentage, 33%, experienced a 57% increase in use, and its cost per use went down by$0.11, or 15%.

Libraries, of course, should not make retention decisions based solely on cost and usage information. Based on this information, each Web of Science search at the University of Connecticut cost $2.17 and each WorldCat search cost $0.37 in fiscal year 2003. Based on its cost per search figure, WorldCat is less expensive. Based on usage, Web of Science is more popular. Obviously these two databases offer significantly different content and cost per use comparisons do not reflect the value of these databases to University of Connecticut faculty and students.

In fact, Townley and Murray, citing guidelines for electronic resource selection developed by Pratt, Flannery, and Perkins[7], assert that, in addition to usage data, qualitative measures such as assessments of content, relevancy, and use of features, should be used to make decisions on

selection and retention of electronic information.[8] In a similar vein, there are journal titles a library will always purchase, regardless of cost per use data, simply because the journal is critically important for a key constituency.

At the January, 2004 meeting of the Association of Research Libraries (ARL) Survey Coordinators, the Executive Director of OhioLink, a large library consortium serving libraries in the state of Ohio, reported that OhioLink analyzes cost and usage data in a number of ways, particularly by publisher. Title and vendor data are sorted by volume of use. Annual cost per search by title and publisher is also calculated. Since OhioLink loads vendor products on its own servers, its usage data is arguably more reliable and comparable than vendor-provided use data. OhioLink has collected data for the last five years, enabling consortial managers to determine if product use is consistent or varied. When OhioLink and other consortia renew their license agreements, they can posture and negotiate based on usage and cost per use management information. The pricing practices of publishers whose total consortial cost per download is significantly higher than other publishers can be questioned, as can vendors whose prices increase faster than usage.

Another consortium member reported at the same ARL meeting that cost per use data was used to cancel specific journal titles after negotiations with a publisher for their package of electronic journals broke down. Faculty who might have previously argued the value of subscribing to all of the journal titles in the package reportedly supported canceling the full package in favor of subscribing to selected electronic journal titles when they were shown some of the cost per use data for titles that were ultimately cancelled.

The Colorado Alliance of Research Libraries is an American consortium of academic and public libraries in two states, Colorado and Wyoming. It offers its members several electronic collections through a statewide database purchase. The Alliance's share of the statewide purchase is calculated based on its share of total use. The state's cost of subscribing to the thirteen databases that comprise the package is roughly $600,000 a year. The databases are then offered to more than 100 libraries, ten of which are Alliance members. Each library in the state is assigned a share of the total cost, based on their annual usage, with a $350 minimum. The ten libraries who are members of the Colorado Alliance of Research Libraries used these thirteen databases approximately 1.3 million times in a recent fiscal year, which represented about 56% of the total state-wide use. Their share of the total cost to subscribe was consequently 56% and it consequently cost each Alliance member about $0.25 for each use of these databases.

As much as librarians may dislike the idea, in our current information world we are co-dependent on publishers and aggregators. Publishers may also take a "cost per use" view of the world and, when they do, it can look like the following publisher-supplied data[9]:

| *Publisher* | *2002 Downloads* | *Average Price* | *Per Use* |
|---|---|---|---|
| | | GBP £ | USD $ |
| Emerald | 3,062,502 | 3.43 | 5.46 |
| IOPP | 3,093,655 | 2.88 | 4.58 |

From this perspective, Emerald and the Institute of Physics Publishing (IOPP) counted use as the download or printing of an article. The price per use from the publisher's perspective was calculated by dividing their annual online revenue (derived from site licenses, consortium deals, pay-per view-traffic and 50 percent of the revenue from their print journals) in 2002 by the total number of downloads. Half of the revenue from print journal subscriptions was included because both publishers offer online access to print journal subscriptions.

Thus, we see that publishers, by considering all related revenues, are taking a more comprehensive approach to average price per use than libraries and consortia typically do when they calculate unit costs based solely on subscription costs.

**Conclusions**

This paper has reviewed some of the pioneering efforts to develop cost per use data for electronic collections. It has noted some of the problems inherent in measuring print and electronic journal usage, whether by title, publisher, search, article download, or some other variable. It has shared some of the analyses that libraries and consortia currently undertake to make and review collection development decisions for the electronic collection comprised of e-journals and databases and has taken a brief look at cost per use data from a publisher's perspective.

A number of trends begin to emerge for library managers to consider. First, as John Cox cautions in the title of a recent article, "Don't Confuse Price With Value." Librarians need to remember that commercial publishers undoubtedly know which of their journals are considered invaluable or indispensable in a given field. For these journals, a relatively inelastic market and pricing structure is at work. At the University of Connecticut, for example, *Brain Research* is our second highest price journal subscription. Its cost per use, however, at $9.12 per download, is exceeded by about two-thirds of the titles we receive through *ScienceDirect.* While the list price of *Brain Research* (about $20,000 annually) remains difficult to grasp, in fact there are many other Elsevier journals at the University of Connecticut that we subscribed to in the print environment that appear to yield less value for the price. In the print environment, price, prestige, and perceived need largely influenced collection decisions. In the electronic environment, librarians may be seeing more tangible evidence of actual value for the price paid.

The University of Connecticut Libraries are also finding that some of the journals we subscribed to in the print environment do not receive as much use as journals we did not subscribe to once we offer broader access to journals through publisher and aggregator packages. Some journals whose prices seemed preposterous in the absence of use data seem more reasonable in the context of cost per use. Other reasonably priced journals, which survived earlier print journal cancellation projects, now seem expendable given their low-recorded use. Granted, we had to be more selective in the print environment than libraries with a greater number of print subscriptions. Given our purchasing history, the data we now have available raises important questions about our previous selection policies.

Different libraries can be expected to have different findings when they employ cost per use statistics. Nevertheless, electronic journal cost per use data, commercial document delivery

charges, and institutional interlibrary loan/document delivery costs provide library managers with cost per use comparisons that were not readily accessible in the print journal environment.

Second, we can anticipate that electronic journal use data will become more reliable. COUNTER released its first Code of Practice in January, 2003. While the Director of HighWire Press feels that stopping at the requirements of COUNTER compliance would fail to provide "useful and/or interesting information that fell outside the COUNTER specification,"[10] COUNTER is important in that it represents an international effort, reinforced by compliance audits, that involves librarians, publishers, subscription agents, and other stakeholders in a coordinated effort. The HighWire Press continues to provide users with "extra-COUNTER information above and beyond COUNTER compliance, but it is also COUNTER compliant.

Lastly, library managers can learn from the pioneering cost study performed at Drexel University. The operating costs in the print environment, even for current year subscriptions, typically exceed the subscription cost. The operational costs for electronic journals on average represented only about 25% of the total cost per use ($0.45 of $1.85) and only full-text database journals (which were used heavily and appeared to be very cost effective) had a higher operational cost per use than subscription cost per use. We can expect lower operating costs in the electronic journal environment than we experienced in the print environment.

Librarians strive to connect users with information they seek to satisfy their information needs. Certainly cost per use data needs to be studied over a period of time to smooth out fluctuations. Decisions should not be made on cost per use data alone. Nevertheless, recent and emerging trends in the utilization of cost per use data to manage electronic collections hold tremendous potential for libraries to increasingly employ improving cost and use data to support collection development and management decisions and better serve our user populations.

**Notes**

[1] www.projectCounter.org

[2] Carol Hanson Montgomery and Donald W. King, "Comparing Library and User Related Costs of Print and Electronic Journal Collections: A First Step Towards a Comprehensive Analysis," *D-Lib Magazine* 8, no. 10 (October 2002) http://www.dlib.org/dlib/october02/montgomery/10montgomery.html.

[3] Oliver Obst, "Patterns and Cost of Printed and Online Journal Usage," *Health Information and Libraries Journal* 20 (2003), 22-32.

[4] Philip M. Davis, "Patterns in Electronic Journal Usage: Challenging the Composition of Geographic Consortia," *College and Research Libraries* 63 no. 6 (November 2002) 484-497.

[5] B. Bauer, "Eine Benutzungserhebung als Entscheidungshilfe fur ein langfristiges Zeitschriftenkonzept. Zeitschriftenevaluierung an der Zentralbibliothek fur Medizin in Wien*,"Mitteilungen der VOB,* 51 no. 2 (1998), 42-52.

[6] Pongracz Sennyey, Gillian D. Ellern, and Nancy Newsome, "Collection Development and a Long-Term Periodical Use Study: Methodology and Implications," *Serials Review* 28 No. 1 (2002) 38-44.

[7] Gregory F. Pratt, Patrick Flannery, and Cassandra L.D. Perkins, "Guidelines for Internet Resource Selection," *College and Research Libraries News* 57 (March 1996), 134-135.

[8] Charles T. Townley and Leigh Murray, Use-Based Criteria for Selecting and Retaining Electronic Information: A Case Study," *Information Technology and Libraries*18 no. 1 (March 1999) 32-39.

[9] John Cox, "Don't Confuse Price With Value – In Academic Publishing, Electronic Is Better," *Against the Grain* 15 No. 6 (December 2003/January 2004) 92-3.

[10] John Sack, "The Beginning of Value Assessment: Usage Information in the E-Journal Age," *Against the Grain* 15 No. 6 (December 2003/January 2004) 36-40.

## Acknowledgements

The author is grateful for the assistance provided by Jim Self at the University of Virginia, George Machovec at the Colorado Alliance of Research Libraries and Deborah Sanford at the University of Connecticut for providing unpublished cost and use data presented in this paper. The author would also like to thank Terry Plum from the Simmons Graduate School of Library and Informatin Science for reviewing the paper and suggesting improvements. Any mistakes, misrepresentations, or other errors are, of course, the author's.