

ONLINE DATABASES
Inundated with Data

By Carol Tenopir

IN THE OLD DAYS, IT WAS DIFFICULT to gather information about the use of library collections. Door counts, circulation records, and tick-marks at the reference desk gave crude approximations. To obtain data about periodical use, patrons were asked to leave used periodicals unshelved, and librarians guessed how many articles were read.

Digital periodicals, ebook collections, and virtual reference inundate libraries with usage data. Some libraries collect it across systems, but most rely on data provided by vendors. If the data allow vendor comparisons, and are consistent and accurate, librarians learn more about their collections.

Enter COUNTER

The purpose of project COUNTER (**C**ounting **O**nline **U**sage of **Ne**T**w**orked **E**lectronic **R**esources) is to standardize data collection and reporting. Usage reports from vendors that belong to COUNTER comply with certain standards. Many libraries already receive the reports or are COUNTER members and can influence the standards. Release 1 of the COUNTER Code of Practice for Journals and Databases came out in 2002; Release 2 will be out in January 2006. Draft Release 1 of the Books and Reference Works standard came out in January 2005.

The COUNTER Journals and Databases standard calls for basic data reports to libraries in Excel spreadsheet format. More detailed reports are recommended but are left to the discretion of the vendor. For journal usage there are two basic required reports: 1) number of successful full-text article requests by month and journal (separated by PDF and HTML in the new standard) and 2) turnaways (instances in which the library collections could not provide the item requested) by month and journal. For database usage there are three basic reports: 1) total searches and sessions by month and database; 2) turnaways by month and database; and 3) total searches and sessions by month and service.

Carol Tenopir was awarded a 2005–07 grant from the U.S. Institute of Museum and Library Services to prepare a cost-benefit model based on usage data for all types of libraries.

The data help identify cyclical patterns, which journals or databases are being used, how the needs of patrons match the e-collection, and whether users prefer HTML or PDF for full-text articles. More interesting, the turnaway counts show where the collection fails, measuring the frustration levels of researchers denied access to items that looked fruitful.

Data must be added

COUNTER-compliant data from major vendors is not enough. It cannot be used without further work. Vendor-by-vendor reports must be combined to compare across vendors. Journal use varies from month to month on regular cyclical patterns, so analyzing a year's worth is more valuable. Raw data only shows how many times a database or journal was opened or downloaded. Each library must add some information about its specific context to make the COUNTER reports meaningful.

> Turnaway counts show where the collection fails, measuring the frustration levels of researchers denied access to items that looked fruitful

Adding the price for each system or database allows libraries to calculate relative cost per use. Since licenses often prohibit price sharing, such cost data can be used only for internal decisions. The data and price may reveal that lower-cost products are not quite the bargain they seemed to be.

Use data and cost per use can be misleading unless the size of the user population is considered. If, for example, a small physics department gets high usage of physics materials relative to the user population, the value of providing access is high per user. Usage log amounts should be weighted and adjusted for the size of each subject population served.

Differences in disciplines

Characteristics of disciplines must be considered when measuring journal downloads. My research on faculty reading patterns with Donald W. King (see web.utk.edu/~tenopir) shows that medical faculty read two to three times as many journal articles on average than humanities or engineering faculty and a bit more than science or social science faculty. These measures of reading have held true over three decades, indicating that they relate to the nature of research in each discipline, not to the value of the reading by faculty. Humanities faculty rely on other sources such as books and primary materials, so while each journal reading may be valuable, there will not be as many. Engineers read many specifications and reports in addition to journal articles.

Consistent usage reports help librarians rethink important collection development decisions and can justify growing e-journal collections. However, the same database or journal may be available from more than one vendor, and each library has a unique environment, so even COUNTER-compliant data requires some adjustment to provide an accurate picture.

Carol Tenopir (ctenopir@utk.edu) is Professor at the School of Information Sciences, University of Tennessee, Knoxville