

# RETHINKING LIBRARY STATISTICS IN A CHANGING ENVIRONMENT

By **Richard W. Boss**

Public libraries have been keeping voluminous library statistics for many decades to focus their resources on needed services, demonstrate the value of their services, aid library directors in the administration of their libraries, and to satisfy the needs of governmental planning bodies. Historically, the emphasis has been on profiling communities and library users, collection sizes and growth, and circulation. Only recently have state library agencies, the primary collectors of data, begun to gather data about the availability of patron access to the Internet, the number of PCs available to patrons, and the electronic resources available. An examination of public library data posted on the Web by a dozen states selected at random suggests that as useful as the data might be for state and federal planning [particularly in response to the reporting requirements of the National Center for Educational Statistics as authorized by the National Education Statistics Act of 1994], and to assure the equitable distribution of grant fund [especially LSTA funds], the data are not adequate for libraries seeking to respond to the needs and expectations of their constituencies and adjust to changing patterns of library use.

## Today's Need

The growth of library collections has slowed and circulation figures are rising only slowly, if at all. While libraries note in their annual reports and budget proposals that they have shifted resources to the purchase of access to electronic products and services, and the use of these by patrons often is a substitute for the circulation of print materials, many directors have not been able to demonstrate conclusively that their libraries' total activity levels have increased.

In addition to a shift from print to electronic sources of information in libraries, there has also been a dramatic increase in access to a library's electronic resources from homes, offices, and schools. Such activity is not measured by the number of visits to the library, therefore, it must be measured at the gateway through which users-in-library and remote-access library resources, whether the resources are located on a library server or the server of an online reference service.

What is needed is a new set of data elements and tools to aggregate the data elements into a single picture of a library's activity.

Library activity includes circulation, interlibrary loan, in-building use, access to electronic information, reference questions regardless of how submitted, and group presentations. It is rare library that cannot document an increase in activity level when all of this activity is documented.

## Required Data Elements

At a minimum, the following data elements should be considered by a public library [Those in italics are not commonly kept today, although it may be possible to determine them by manipulating available data]:

### Population and Registered Borrowers

- Size of population served.
- Income level of population served.
- Non-English speaking population served (by language).
- Number of registered borrowers from the jurisdiction served.
- Percentage of population in the jurisdiction served who are registered borrowers.
- Number of registered borrowers from other jurisdictions, including reciprocal borrowers, fee borrowers, guest borrowers, etc.
- Number of registered borrowers by zip code.
- Number of registered borrowers by language preference.
- Number of registered borrowers by age level.
- Number of registered borrowers by sex.

### Budget

- Total budget from all sources.
- Total budget by source of income.
- Expenditure per person in population of jurisdiction served.
- Expenditure per registered borrower.
- Hours open per week (aggregate of all outlets).
- Expenditures for staff, including benefits.
- Number of staff per 1,000 registered borrowers.
- Expenditures for library materials, including print, A-V, and microform.
- Expenditures for library materials per registered borrower.
- Expenditures for electronic products and services.
- Expenditures for electronic products and services per registered borrower.
- Expenditure for all library materials and electronic products and services per registered borrower.
- Value of electronic products and services available as the result of consortia and/or state agency purchases, but not reflected in the library budget.

### Facilities

- Expenditures for facilities, including technology and other equipment.
- Outlets per 1,000 registered borrowers.
- Number of PCs available to patrons.
- Number of PCs available per 1,000 registered borrowers.

## Collections

- Number of print items in the collection.
- Number of print items per registered borrower.
- Number of non-print items in the collection, including A-V, microform, and electronic.
- Number of non-print items per registered borrower.
- Number of print serials available.
- Number of print serials per registered borrower.
- Number of electronic serials available.
- Number of electronic serials per registered borrower.
- Total number of unique serials available.
- Number of unique serials per registered borrower.
- Number of electronic databases available.

## Usage

- Annual Number of visits to library.
- Average number of visits per population served.
- Average number of visits per registered borrower.
- Circulation of locally held print materials.
- Circulation of locally held non-print materials, including A-V, microform, and electronic.
- Circulation to homebound in above totals.
- Circulation per registered borrower.
- Number of items loaned to other libraries, including photocopies.
- Number of items borrowed from other libraries, including photocopies.
- Turnover rate of circulating materials (number of circulations divided by the number of items in the collection)
- Number of in-library uses of print materials (reshelving count, minus circulation count).
- Number of in-library uses of non-print materials (reshelving count, minus circulation count).
- Number of patron access catalog searches.
- Number of patron access catalog searches per registered borrower.
- Number of accesses to electronic products and services per registered borrower.
- Number of page views of electronic products and services.
- Number of page views of electronic products and services per registered borrower.
- Number of accesses through the gateway or portal to all external resources.
- Number of accesses through the gateway or portal to all external resources per registered borrower.
- Number of access to the library's Web site.
- Number of visits to key pages on the library's Web site, including number of sessions and page views.
- Number of in-person reference questions.

- Number of e-mail, telephone, and mail reference questions.
- Total number of reference questions.
- Total number of reference questions per registered borrower.
- Number of presentations to groups.
- Number of participants in group presentations.
- Total number of transactions or contacts represented in all of the foregoing (i.e., activity level).
- Transactions or contacts per registered borrower.
- Average transactions or contacts per hour open.
- Average cost per transaction or contact.

A library may choose to base its ratios on the size of the population, rather than on the number of registered borrowers. As long as the comparison from year to year uses the same basis, it will be possible to accurately determine the change in activity level.

## Collecting Data

The majority of the data can be collected using the standard statistical package and report generator of an automated library system. Unfortunately, while all automated library systems keep statistics on acquisitions, cataloging, and circulation activity, some do not keep track of the number of searches against the patron access catalog because that is a resource-intensive activity. It is possible to include this capability as an extra-cost option. The cost consists of minor programming and, possibly, some hardware.

With rare exceptions, if data is retained in the automated library system, the system's report generator can be used to create a report that meets the specific needs of a library. The data can also be downloaded to a database or spreadsheet on another hardware platform.

A major change in data gathering that most libraries need to make is to measure the amount of in-house use that library materials receive. To obtain this number, staff must scan all materials left on reading tables and used to answer reference questions before reshelving. This is not only a good activity measure, but can be an aid in collection development. Materials that are never taken off the shelves may be candidates for storage or withdrawal.

Interlibrary loan statistics can be kept by the automated library system if the system's ILL module is implemented. The relatively inexpensive module keeps track not only of the number of items lent and borrowed, but also copyright compliance and the "balance of trade" with other libraries. Alternatively, the ILL unit can keep track of its statistics and submit them periodically for entry into a database or spreadsheet.

The implementation of portal technology gives a library the opportunity to gather even more data electronically. Most portals include a statistical package. MuseGlobal's software, which is used by five automated library system vendors, is particularly good for capturing data on all uses of the portal to access electronic resources, including the local

patron access catalog, locally stored electronic resources, the patron access catalogs of other libraries, online reference services to which the library subscribes, and URLs to which a library has provided links.

For libraries that don't implement portal technology, access control software on each PC or on a server can measure use of electronic resources. Almost all such products—even those designed to block sites deemed unacceptable or to filter content—recognize and authenticate users and maintain statistics on accesses to desktop applications, online reference services, specific URLs, and the Internet as a whole. In the case of Comprise's Internet Manager and CybraryN's Internet Control, the blocking and filtering can be turned off, and only the access management and tracking activated. EnvisionWare's Launch Command has the management of access and recording of activity as its primary focus.

There is an effort underway to develop a single, international extendible Code of Practice that allows the usage of online information services to be measured in a credible, consistent, and compatible way using vendor-generated data. The COUNTER Code of Practice specifies: the data elements to be measured, definitions of these data elements, usage report content, format, frequency and methods of delivery, protocols for combining usage reports from direct use and from use via intermediaries. Release 1 of the COUNTER Code of Practice issued in December of 2002 focuses on the usage of journals and databases. A number of database producers and Ebsco, a major aggregator of databases, have committed to adopt the Code of Practice, although it will be difficult for an aggregator to measure use when the database is not mounted on its servers. It may have to rely on the database provider for the data. Future releases of the Code will extend the scope of COUNTER not only to other content types, but also to more detailed levels of reporting for each content type. Release 2 will definitely specify that the individual article in a journal be identified. It may be another two years before Release 2 becomes available. See [www.projectCOUNTER.org](http://www.projectCOUNTER.org) for additional information.

## **Manipulating and Presenting Data**

A simple database or spreadsheet on a PC is all that most libraries need to maintain and manipulate their data. A *database* is a collection of information that is organized into fields (a single piece of information), records (a complete set of fields), and files (a collection of records). One can download information from an automated library system into any one of several databases. Among the most widely used are Microsoft Access, Lotus Notes, and Corel Paradox. Access and Paradox are parts of office suites that are widely used in libraries. All of these database products provide templates for entering information and simple tabbed windows for designing tables.

A *spreadsheet* is a rectangular grid used to collect and calculate data. It is basically an electronic version of a ledger book. It allows one to figure totals, calculate formulae, sort information, and translate the information in a spreadsheet into graphs. It can also be used as a database manager to keep track of information that can be held in a list. A spreadsheet is a better tool for manipulating data than a database, especially when one

wants to change one data element that may affect other data elements. For example, one could ask how the staff-to-registered-borrower ratio would change if the population increased by ten percent, but the number of staff remained unchanged; and how many staff would have to be added to maintain the same ratio. Among the most widely used spreadsheets are Microsoft Excel and Corel Quattro. Each is part of an office suite widely used in libraries.

## Sources of Information

The best single source on methodology for gathering and using library statistics is Arthur W. Hafner's *Descriptive Statistic Techniques for Librarians*, Chicago: American Library Association, 1998. While it does not recommend data that should be gathered, it discusses basic methods for data analysis, data comparisons in the library environment, presenting numerical data in tables and charts.

Also useful is G. E. Gorman's "Collecting Data Sensibly in Information Settings," Written in 1999, it is available as ERIC document ED 441441.

Among the many publications on measuring the use of electronic publications and services, particularly useful is P.R. Young's "Measurement of Electronic Services in Libraries: Statistics for the Digital Age," *IFLA Journal*, volume 24, number 3, 1998, pages 157-160.

A newly released American Library Association publication by Diane Mayo and Jeanne Goodrich entitled *Staffing for Results: A Guide to Working Smarter* provides detailed guidance for documenting staff workloads and allocating staff.

©2006 American Library Association. [Copyright statement.](#)