MEASURING IN A VIRTUAL ENVIRONMENT

This issue of *Library Technology Reports* explores the topic of measuring the use of electronic content and services provided by libraries. The approach taken is a practical one that helps librarians think about the issues involved and learn some practices to effectively document how library users take advantage of electronic content and services.

Balance and perspective

Anyone involved with libraries notices the growth in volume of electronic or digital information forms. But by no means are libraries becoming completely electronic. A realistic assessment reveals that most libraries continue to serve their traditional roles, complemented with an increasing amount of electronic components.

This report focuses on measuring the use of the electronic and digital services and content offered by libraries. You will learn how to measure each aspect of the digital side of the library, but take care to use these measures in the context of the broader scope of the library's services, giving the traditional services equal voice. For each aspect of the library that takes some electronic form, there are likely traditional services that should be given equal statistical attention.

Picture a continuum with the traditional library on one end and the theoretical, all-digital library on the other. Few, if any, libraries today are on either end of the continuum. Most libraries continue to have physical buildings, large collections of books, and patrons who come inside to read and do research. Patrons talk face-to-face with a librarian when they need guidance in finding materials. On the other end of the spectrum, an all-virtual library might exist without a building, offering a collection of all electronic materials that can be used from the convenience of the user's home or office. The collection would include a vast array of electronic books, would have the full text of journals and periodicals online, and would provide the assistance of a librarian through a live online reference service.

For each aspect of the electronic components you seek to measure, the results should be placed into perspective, balanced by the larger context. Although a given statistic about an electronic library service may be interesting and impressive considered by itself, how does it compare with similar activities conducted by the library in nonelectronic form? Measure both the electronic and traditional forms of the service and see how the library stands in the continuum for this aspect of its operation.

This report considers ways to measure each of the major aspects of the library that have electronic components, including the online catalog, the library's Web site, electronic content procured from vendors, and digital collections.

Throughout this report, the terms *electronic, digital,* and *virtual* carry the same meaning. These terms all describe computer technology. No nuances of meaning are intended when using these terms. Likewise, the words *use* and *usage* have the same meaning.

Increasing electronic investments

The significant investment in digital content and electronic services make measuring their use important. Librarians of all types devote much of their time and resources to the creation, purchase, or maintenance of electronic resources and deliver an increasing amount of services via the Web to remote patrons. The level of investment in electronic resources and services by libraries is likely to continue to increase. Just as libraries have previously developed policies, systems, and procedures toward the measurement, assessment, and management of traditional library materials and services, the same processes need to take place in the electronic arena.

The way libraries measure and assess their services must adapt to the changes that take place in patterns of use. The increased amount of electronic content naturally leads to higher levels of use beyond the physical confines of the library. Counting activities that take place inside the library are easier than measuring how remote patrons use the library's electronic resources. If you expect a trend toward increased use of electronic resources and services by remote users, implementing ways to measure their use and to assess their effectiveness becomes critical.

According to a recent Association of Research Libraries' (ARL) Report (ARL Supplementary Statistics, 1999-2000, p. 5), 12.9% of the library materials budget for ARL's large academic and research libraries goes toward electronic resources, and that spending is rapidly increasing.

Measurement and assessment have gained renewed attention as important themes in libraries. Many library organizations have initiatives underway, including ARL, International Coalition of Library Consortia (ICOLC), Digital Library Federation, National Information Standards Organization (NISO), and the U.S. National Commission on Libraries and Information Science.

The measuring of the use of electronic resources and services complements other related tasks librarians are likely pursuing in other aspects of their operations and collections. These efforts form an important part of a library's overall strategy to measure its impact on its user community. Measurement and assessment of electronic resources and services document the library's accomplishments and provide information that can be used to make any needed improvements in quantity or quality. Public libraries have more modest investments in electronic resources, averaging about 1% of overall budgets on purchased electronic titles, with another 3% put toward providing access to electronic content by subscription.

Measuring the use of electronic resources itself can be a time-consuming activity. Measurement may involve the purchase, implementation, and operation of additional software, and negotiations with vendors to provide adequate use statistics. Most of the work requires staff time to gather quantitative data from multiple sources and normalize them into standard form to analyze and interpret the results.

No doubt, efforts in measurement and assessment add to the overhead of maintaining electronic assets. Like all other support activities, efficiency demands the least amount of resources be devoted to this set of tasks while still achieving desired results. The process of creating statistics often requires a disproportionate level of time and effort. A balanced and sustainable effort yields better results than an overly ambitious one that exceeds the time and energy available from the library staff.

Audience of this report

This report targets a broad range of administrators and practitioners in libraries and other information-oriented organizations. Library directors, unit managers, and other decision makers may use this report to find information regarding the types of measures related to electronic resources that are available and to provide some perspective on their relevance to management decisions. The practical, how-to information in this report assists Web site designers and implementers in gathering statistics and in optimizing systems to produce adequate statistics. Electronic collection development staff benefit from the sections that show how to measure the use of individual resources.

Readers should have general familiarity with the concepts and issues related to electronic resources in the typical library Web environment. Even though this report covers some technologies that might be employed to collect and measure use, previous technical knowledge is not necessary. Any technical details are explained. Some material specifically for programmers and technical staff is included. These sections are identified and are not essential for understanding the report as a whole.

Scope

This issue of *Library Technology Reports* covers the following topics:

- Statistics related to the integrated library system (ILS)
- Measuring and evaluating the use of the library's Web site
- Electronic reference services, including e-mail, chat, and Web-based systems
- Vendor-provided use statistics
- Locally generated electronic content for digital library systems

The elusive virtual library user

When organizations conduct business on the Web, building customer relationships is key. In the retail arena, for example, a direct relationship exists where a customer purchases a product from a company. Money changes hands, and the customer's satisfaction can be measured. But the relationship a library has with its virtual users isn't as direct. The library and the user may have a low level of awareness of each other. People may take advantage of resources acquired on their behalf and made available to them by the library, with no knowledge of how that access was made possible.

Due to the lack of opportunities to brand a particular library's information resources, many users consider library-provided resources as just part of the Web. Likewise, libraries often have only indirect knowledge of how the resources they procure—at considerable expense—are used.

The act of reading or printing a page of information on the Web is a less well-defined transaction but is a core service of the library. As libraries strengthen their relationships with their online customers, the number of measurable service transactions should increase. A growing number of library automation systems with personalization features allow a library user

to log into an online public access catalog (OPAC) or portal environment to receive a set of customized services. These personalized environments empower the library with an increased ability to measure how specific communities use electronic resources and services.

Virtual visits

In the last few years many libraries have experienced major changes in the ways their patrons use their facilities. In a generalized way, most libraries see a moderate shift in the continuum from traditional toward virtual library use. Many libraries report declines in building entrance counts, number of questions asked at reference desks, and numbers of items checked out. But the level of activity for a library's Web-based facilities is increasing. To substantiate this trend and provide background and perspective for the changing patterns of library use, you need to develop a set of comparative statistics describing what might be called virtual visits to the library's Web-based resources versus physical visits to its buildings.

Making measurement and assessment practical

Gathering statistics helps support specific needs of library managers and administrators related to processes of decision-making and resource allocation. With traditional services and physical materials, many libraries have routines for gathering detailed statistics. Yet, reexamination of processes may reveal some of these statistics are superfluous, having no specific relevance to assessment or management decisions of resource allocation. The level of effort put into measuring user activity must be balanced. The basic question is: how much time do you want to devote to measuring what you do at the expense of doing less? The same balance is needed with the electronic world. Although most libraries are working up to a level of adequate measurement and assessment, they also must not create an onerous level of overhead that takes too much away from the overall effort.

Some of the obvious applications of statistical information related to electronic collections and services might include the following topics.

Budget justification to external funding sources

Libraries seldom find themselves in a position where they do not have to justify their levels of funding and how those funds are dispersed. As increased proportions of financial resources go toward electronic efforts at the expense of traditional materials and services, external funding agencies want to have evidence of the soundness of this shift.

Strategic planning

An understanding of the level of use of electronic resources and services is a key input into the strategic planning process. Decisions on the relative balance between traditional and digital library services and whether to embark on new electronic initiatives often depend on the performance of the library's current efforts.

Resource allocation within the library

Library decision makers are a main consumer of statistics. Library directors, unit managers, and others in positions with the authority to allocate resources rely on data regarding the relative importance of competing interests. A detailed understanding of the levels of use of electronic resources is vital to ensure appropriate levels of funding are provided in the budget planning process.

Collection development

On a more granular level, understanding the level of interest for each electronic resource is important. Much electronic content takes the form of a renewable subscription and is not a once-and-for-all purchase. For better or worse, electronic content tends to be licensed, not bought outright. Although initial procurements of electronic resources must be made based on anticipated use, decisions regarding the continuations of subscriptions can be made based on measured prior use.

The standard selection and deselection processes related to collection development apply to the portions of the library's collection that are electronic. The stakes are often high with electronic content. The initial investment, especially for aggregated products, can be high, so the selection process must be made carefully in relation to the overall collection funds available.

Careful thought must also be given to what electronic subscriptions to continue each year. Some electronic content can be acquired though a one-time purchase that entitles the library to permanent access of the material. These titles carry little cost of ongoing ownership. Depending on whether the content is held on local servers or is remotely hosted, the only direct costs may be related to the prorated costs of maintaining the equipment. Titles that involve an annual subscription have significant ongoing costs, which must be assessed periodically to ensure their cost level is in line with the amount of use and available funding.

Don't base decisions only on sheer volume of use. Many electronic titles offer content over a broad slate of interests, and the volume of use may well be an appropriate measure of value. Some electronic resources are highly specialized and target a more narrow range of users but still carry high value. The point is not to take too simple a view of the statistics, but to infer the more complex relationships of use by specific user communities.

Use statistics can help derive a relative value for each resource. Determining the unit cost per instance of use for electronic resources is often helpful. Armed with knowledge of the total cost of a resource and its respective use data, the cost per view can be determined. Resources with a high volume of use may have low costs per view, even if they carry a high subscription cost.

Usability improvement

Counting the number of times an item of electronic content is accessed may not only measure interest but may also be an indicator of usability. Resources that are difficult to find within a library's Web environment, or contain content whose nature is not well-described, achieve levels of use

below their value and potential. In this regard, use statistics serve as data in usability analysis. When actual use falls below projected interest, further work may need to be done on the organization and presentation of the library's Web site.

Level of detail

The effectiveness of statistics often depends on the form in which statistical data are presented. Raw data are rarely useful and must be organized and summarized before they can be analyzed and interpreted. The level of appropriate summarization varies according to the intended audience. People who are working on day-to-day decisions regarding electronic resources often require fairly detailed information. High-level library administrators likely prefer broader views of the data, but with enough detail to understand how the summaries were derived. Statistical reports generated for people outside the library administration are likely the most generalized.

Longitudinal data

Although the current levels of use for any given piece of electronic material is useful to know, studying how use patterns change over time is more interesting. Building a set of longitudinal data helps in this regard. Keeping use counts in each category over a long period of time provides the ability to study interest levels for any title, as well as shows broad trends regarding the use of electronic materials.

Data mining

Data mining involves the ability to continually extract information from a data set according to ongoing needs and requirements. The process involves analyzing data from various perspectives and looking for new patterns. You can drill down through a data set to extract information of interest as you pass through each layer or category of data. Although data mining is often associated with large, relational databases, it can be applied to any set of organized information. The data collected from a library's use statistics-gathering efforts can become part of a data store that can be exploited with data-mining techniques.

Multiple sources

Given most libraries' involvement in many types of electronic resources and services, expecting that any single system or source provides a comprehensive framework for gathering use data is unrealistic. Multiple ways for users to access library-provided resources always exist. Providing multiple access points increases the opportunities for patrons to find the resources that are useful to their interests. In a typical library environment, a given resource may be offered through the OPAC, through other Web-based finding aids, or in subject-specific resource guides.

Even when the initial use of an electronic resource takes place through a library's Web site, subsequent uses may not. Users may either remember the

URL or set a bookmark in their Web browser for resources they find especially useful.

Some of the sources a library needs to rely on for a complete accounting of the use of its electronic resources include:

- Access logs from the library's Web server
- Use statistics supplied by vendors and service providers
- Logs generated by the library's proxy server
- Logs and reports from any locally implemented journal locator services
- Logs and reports from any locally implemented digital library products
- ILS reports: circulation reports, transaction logs, ad-hoc reports.

Problems with unmediated use

One of the major issues in measuring the use of electronic resources involves cases where the resource is provided by the library, but the users access the resource directly from the provider. The library cannot capture all instances of access to its resources. A significant amount of use may occur directly between the end user and the information provider with no intervention by a library server.

Some library users may discover and use library-provided resources without ever visiting the library's Web site. Students, using library computers, often perform a search on an Internet search engine and find information on a vendor's Web site. The student is authorized for access to the resource by virtue of a library-provided subscription. But since the student did not go through the library's OPAC or Web pages, that use can't be counted by any library-side process. Only vendor-supplied statistics can capture this category of use.

Users may also gain their initial access through the library's OPAC or journal finder system, but their subsequent access is accomplished through a bookmark they set in their Web browser.

Given the likelihood that unmediated access will increase over time, the need for providers of electronic services to be able to provide complete use statistics grows. The vendors—in the role of service provider—have better technical ability to measure the use of their systems than the library does in its role as the buyer of those information products.