

## ◆ Following Your Readers with Web Statistics

Presenter:

**Julia Blixrud**

Association of Research Libraries  
Washington, DC

Reporter:

**Thomas Gegeny**

The Center for AIDS  
Houston, Texas

Several years ago, when the World Wide Web was relatively new, counting “hits” or Web-site visits seemed a good way to track Web-site use. However, new types of detailed data are needed as Web-site content grows more complex and sophisticated. For journal publishers, one important point is that Web statistics do not always indicate journal use online. Julia Blixrud, director of information services at the Association of Research Libraries (ARL), reviewed these concepts and explained what is being done by organizations like ARL in attempts to standardize Web statistics.

What can Web statistics be used for? Such information can be used for budgets (justifying expenditures and demonstrating accountability), strategic and operational decisions, and service improvement. To be useful, these data must be *detailed*, *consistent*, and *comparable*. The detail should include use data (titles accessed and articles viewed vs printed or downloaded), user data (repeated questions or searches, such behaviors as viewing tables of contents or references), and failure rates (interface

problems, browser incompatibilities, and user errors). Consistency is important so that there can be standard definitions and interpretations of data. For instance, what defines a search can depend on how a Web site’s technical platform is set up. Frequent, clear, and consistent reports from vendors are also needed. Data must be comparable within an institution or even between institutions for purposes of benchmarking and measuring performance.

The world of online information has many participants: book and journal publishers, aggregators (vendors who offer several journals as packages), database and e-reference companies, and libraries. Several national and international efforts are under way to create standards and guidelines for measuring and reporting Web statistics. One such effort was the creation of a set of guidelines by the International Coalition of Library Consortia in November 1998; these are to be reviewed and revised in September 2001. They are available on the Web: [www.library.yale.edu/consortia/webstats.html](http://www.library.yale.edu/consortia/webstats.html).

Blixrud also reviewed ARL’s *E-metrics Project*, which involves 24 libraries engaged in a collaboration to assess current practices of statistics gathering and eventually propose a set of standard, tested measures. An initial survey indicates that many libraries collect some data on accessible resources and their related costs. However, the survey found that libraries have very few use data and depend almost solely on vendor reports for such information. In addition, an analysis of vendor reports showed that

not all vendors provide their customers (libraries) with any statistics and that the types of statistics reported, report delivery, and report format are inconsistent.

Another important issue is user privacy (for instance, the use of cookies to track user information, which users can block from being placed on their computers). A user might not mind if the amount of time on a page is tracked, but might mind if the topic he or she reads is tracked. Privacy statements explaining what statistics will be tracked and how they will be used might be a good idea. Other concerns related to Web statistics include the absence of an established baseline or history, a lack of standards and definitions, and the undetermined value of use data—what does it mean to access a table of contents or a specific article?

Obviously, Web statistics are of great significance to publishers, libraries, and users alike. Meaningful statistics must combine use data with user data. Put simply, an electronic resource that receives few hits but provides substantial information to visitors is not necessarily less valuable than a resource that receives many hits but does not provide visitors with the information they need. The efforts described above and others are paving the way for a better understanding of these issues and underscore the importance of standardizing how we track a technology that has revolutionized the publishing industry. 