Why and How to Measure the Use of Electronic Resources

Jean Bernon

Directeur du Service commun de documentation,
Université Jean-Moulin Lyon 3, 6 cours A. Thomas,
69371 Lyon Cedex 08, France,
bernon@univ-lyon3.fr

Abstract

A complete overview of library activity implies a complete and reliable measurement of the use of both electronic resources and printed materials. This measurement is based on three sets of definitions: document types, use types and user types. There is a common model of definitions for printed materials, but a lot of questions and technical issues remain for electronic resources. In 2006 a French national working group studied these questions. It relied on the COUNTER standard, but found it insufficient and pointed out the need for local tools such as web markers and deep analysis of proxy logs. Within the French national consortium COUPERIN, a new working group is testing ERMS, SUSHI standards, Shibboleth authentication, along with COUNTER standards, to improve the counting of the electronic resources use. At this stage this counting is insufficient and its improvement will be a European challenge for the future.

Key Words: information resources; use studies; library statistics; France; COUNTER

In 2006 the Library Department of the French Ministry for Higher Education (SDBIS) asked the Director of the National Bibliographic Agency (ABES), Ms Sabine Barral, for a report on the national counting of electronic resources use. A group of library directors was set up to debate with her at each step of the study. The report was completed at the beginning of 2007. I participated in the working group and the following remarks are mainly extracted from the discussions of the group and from the report.

Knowledge of library activity is the main aim of measuring the use of electronic resources. We need at least the same level of knowledge for electronic resources as for the printed collections. The knowledge of the activity for printed collections is based on the fact that different libraries have a common model of description with three sets of definitions: definition of the document types (books, serials, maps etc.), definition of the use types (loan, consultation, photocopying), definition of the user types (undergraduate, master, doctorate, researchers, teachers, faculty members, external users). These sets of definitions are not very clear for electronic collections and the working group tried to fix them.

The following *document types* were adopted: bibliographic databases, full-text databases, serials, databases mixing data and serials, e-books and monographs, websites. The group underlined that the electronic activity of the library was not complete with just the publisher journals and databases on which the COUNTER model is focusing. The use of online catalogues, of local portals, of open archive documents, of digitised documents and of online reference services should also be measured to get a full overview.

The definition of the *use types* is more difficult because of their great diversity: connection, session, hit, direct search, access through search engine*s*, access through syndication, browsing, displaying, printing, downloading of different formats. Librarians are often puzzled over the exact type of use counted by different publishers, making comparison hazardous. The group followed the COUNTER definitions, keeping three main uses to be counted: sessions and turnaways (rejected sessions), database searches, downloaded documents (serial articles or book chapters). The counting reports must be both monthly and annual for every use. They must detail services when one database offers several services and detail journals when a service gives access to several journals.

Knowledge of the *user categories* of the different electronic resources is the last important issue in measurement of the library activity, and it remains the most difficult one. Users are seen through technical network devices. The IP addresses can be private or public, static or dynamic and cannot be easily matched to user groups. Access through a directory is basically more efficient, but the different technologies (LDAP, CAS, Shibboleth) do not always give the same user information; they are useless for anonymous access to free resources and above all the user information is controlled at the customer level and should not be in any way controlled by the supplier.

Jean Bernon

A complete overview of electronic activity will help the library to *compare efficiently the cost of the resources with their use*. Librarians need to know the ratio cost per use for each database, for each serial title, for each article. In France from 2001 to 2005 the national expenditure on electronic documentation increased fourteen-fold. Science, technology and medicine libraries have been cancelling subscriptions for several years, law and management libraries are now encountering the same difficulties and electronic use is growing quickly in literature and social sciences fields. Libraries have more and more important subscription decisions to make and need to compare the services of different publishers, to compare between publishers and open archives, to compare between printed materials and on-line materials.

The complete overview of the electronic activity is not the same at the different decision levels. The library sometimes needs detailed counts of one specific database for technical checks or quality survey. The library director and university management need some help to choose between different specific resources and to stay within the budget. The regional, national and international decision levels are focused on area coverage and global counts; they first need a reliable comparison between universities.

Which Combination of Methods Can We Use for Counting this Use?

There are publisher tools and university tools. Most of the counts are produced by publishers. But they are very incomplete or inconsistent. Some publishers are supplying no count or useless counts. Some others (especially small or local publishers in the literature and social sciences fields) are using specific tools and the librarians have to make their counts consistent. A minority of publishers are COUNTER-compliant, but these are the international publishers and the biggest library suppliers.

Today universities cannot get a complete overview of the use of their electronic resources and some of them build local tools. This can be a simple web counting of the pages and links to the databases used in the local portal. For example, some library portals are using the Xiti service. This method is useful to get a global measurement and a general comparison. It can give user

information when the portal authenticates them. However, as with every university method, part of the use escapes measurement when users have direct access to databases or access through other portals. More specifically, this method does not give any detail about the use types.

Therefore few universities have experimented with another local method, the *deep analysis of proxy logs.* The results of these experiments are clearly more consistent and detailed than with other methods. They are nearly complete, except for access outside the proxy, but they are using a knowledge database which needs heavy support. Many universities have no proxy or cannot dedicate it to documentation and the network must be set up so that most of the database accesses pass through the proxy. Ms Barral's report included an extensive study of this tool and suggested the implementation of a national release of the knowledge database, but this programme was considered too heavy by the majority of the working group.

Anyway it is evident 1) that electronic resource suppliers can only measure use of their resources, and 2) that only universities can maintain an electronic directory of their members; more and more universities are using centralised authentication. The best method to build a complete and reliable overview of the use will mix publisher/supplier tools with university/customes tools. The new tools are basically mixing the two sides. The Shibboleth authentication system is based on a network of 'identity suppliers' and 'service suppliers'. The SUSHI standard is a complement of the COUNTER standard to be used by local Electronic Resource Management Systems (ERMS). The technical department of the national French consortium COUPERIN is now working on such mixed tools and has set up a working group on ERMS.

Are We at This Stage Able to Exploit the Results and to Compare Them between Institutions?

It is obvious from the previous remarks that the counts of electronic resources use are not reliable enough to be compared within one institution and *a fortiori* between different institutions. Counts of electronic resources are always increasing year on year in every library and every country. It means that the use is increasing, but we do not know by how much exactly. An unknown part of the increase comes from the improvement of the counting tools themselves.

Some figures can illustrate this statement. In the 2003 French national survey on electronic resources, between 10% and 57% of the universities responded to the different questions and the answer rate is increasing year on year. Another survey in August 2006 (to which 64 of 104 French university libraries replied) showed that, in a list of 144 electronic resource suppliers, 36 had no counting system, 37 were COUNTER-compliant and the remaining had very specific counting systems. The 2007 report of my university (Jean-Moulin Lyon 3) lists 42 database subscriptions. Five of them are supplied without count and two of them with a global counting at consortium level without institution level counting. Only two vendors and 6 of 42 subscriptions are COUNTER-compliant.

The conclusion must be that despite the growing use of COUNTER, measuring the use of electronic resources remains at this stage very incomplete and unreliable. However, some progress can be noted especially with regard to measuring tools, and it is an important aim for libraries to reach a reliable standard of measuring in the next few years. This improvement is not only a technical matter. It requires agreements between libraries and between electronic resource suppliers and libraries. It requires common work on standardisation, tests and cross-checks. This challenge has to be taken up urgently. The exchange of information and best practice between LIBER members represents an important contribution to this work.

Websites Referred to in the Text

Barral, Sabine (2007), Mission: Indicateurs d'usages des ressources électroniques, rapport final, http://www.sup.adc.education.fr/bib/Acti/Electro/accueil. htm#mission

COUPERIN working group on ERMS, http://www.couperin.org/rubrique. php3?id_rubrique=59

Xiti service, http://www.xiti.com