Performance Indicators for the Digital Library[1]

by ROSWITHA POLL

The purpose of performance indicators is „to assess the quality and effectiveness of services provided by a library ... and to assess the efficiency of resources allocated by the library to such services."[2]

In the last decades, libraries have developed sets of performance indicators for evaluating the quality of their services. Such indicators have been presented in handbooks and in an international standard. They were for the most part restricted to the traditional library services (collection, lending and ILL services, reading rooms, reference desk).

For the fast developing electronic library services new and special indicators are needed in order to assess the effectiveness of services and to justify expenditure and resource allocation to that sector. In several projects on a national and international scale, such indicators have been developed and tested.

THE PROJECTS

In the first place, statistical data for electronic and web-based services had to be defined and described:

- ISO/DIS 2789 (2000) Information and documentation - International Library Statistics

  The existing standard has been revised and enlarged in order to include statistics for electronic library services. The standard will pass through its last voting process in 2001.

- International Coalition of Library Consortia (ICOLC). Guidelines for Statistical Measures of Usage of Web-based Indexed, Abstracted, and Full Text Resources. 1998

  The guidelines are at present under revision. They define statistics for databases and electronic journals in consortial agreements.

Other projects take the broader view of statistical data and performance measurement for electronic services:

- John C. Bertot, Charles R. McClure, Joe Ryan: Statistics and Performance Measures for Public Library Networked Services. Chicago: American Library Association 2001.
  The project is finished. It tested statistics and performance indicators for evaluating electronic services in public libraries.

- ARL „New Measures Programm" - <http://www.arl.org/stats/newmeas/newmeas.html>
  The project has started in 2000. It aims at establishing usage measures for electronic information resources in ARL libraries.

- EQUINOX. Library Performance Measurement and Quality Management System. - <http://equinox.dcu.ie/>
  The project finished in 2000. It developed and tested performance measures for electronic services in all types of libraries.

- ISO TC 46/SC 8/WG 4: Technical Report 20983: Information and Documentation - Performance Indicators for Electronic Library Services (Working Draft).
  The Working Group 4 is responsible for the maintenance of ISO 11620: Performance Indicators for Libraries. Because of the quickly changing structure of electronic library services, performance indicators for this sector will not yet be presented in a standard, but in a Technical Report.

THE DATA TO BE COLLECTED

When libraries evaluate their electronic services, the goals are similar to the evaluation of traditional services. The main questions are:

- Does the library give access to the information resources users need?

- Do users get access directly and easily?

- Is the equipment for the use of electronic services sufficient and functioning?

- Do the services reach the population?

In order to answer such questions, libraries need to collect data about

- the size of their electronic collection (journals, databases, single documents)

- the use of their electronic services (electronic collection, OPAC, the library's website, document delivery services etc.)

- the expenditure on electronic services (collection, equipment, other).

DEFINITION OF DATA

The most important electronic library service, as in the traditional library, is the library collection, the information sources the library offers to its clientele. But while a collection of print documents could be easily marked off from documents outside the library, for an electronic collection the boundaries are hazy. Libraries can subscribe to documents on remote servers, such licences may be restricted to a certain time period, or libraries pay for a certain amount of accesses to documents for their users (pay-per-view). What then should be included in the term „collection"? Libraries have to define this, as they must justify the expenditure for information resources.

The draft of the ISO statistics standard 2789 gives a new definition of „library collection":

All documents provided by a library for its users

- Comprises documents held locally and documents on remote resources for which access rights have been acquired, at least for a certain period of time

- Access rights may be acquired by the library itself, a consortium and/or external funding

- Acquisition = deliberate selection + securing access rights + inclusion in OPAC or other database of the library

- Does not include simple links to Internet resources.

This definition restricts the term „collection" to documents that have been deliberately selected and catalogued in order to meet the needs of a defined clientele, and for which the right of access has been arranged for the same clientele, in most cases by contractual agreement and payment.

In the projects afore-named, the most difficult task proved to be defining „use" of electronic services. Several possibilities of using electronic services were identified:

- getting access to a service (e. g . by a log-in); such access has also been defined as a „session" or a „visit"

- doing a search, putting a query

- viewing a document or record

- downloading a document or record.

Certain problems occured frequently when the libraries involved in the projects tried to count cases of use:

- The data have to be collected from different sources (the library's servers, consortial servers, suppliers' servers etc.) and thus might be varying or incomplete.

- Search engines can affect the counts.

- Downloading documents from external servers could in most cases not be counted, as the downloading procedure is performed on a local browser.

- The definition of one search or query proved difficult. Should one count one or several searches, if a user modifies the search after getting to many or too few „hits"? Also, an established connection to an electronic document might be used by several users one after the other, so that different cases of use could not be identified.

At the present stage, it seems reasonable to restrict counts of use to the following cases[3].

Session (access, visit)
Established connection to an electronic service, usually a log-in.

Rejected session
Unsuccessful attempt to connect to an electronic service because of requests exceeding the simultaneous user limit.

Document viewed
Full text uploading of an electronic resource, or catalogue record/database entry fully displayed.

When counting use, libraries will probably want to distinguish between use by the own clientele and other users, as they have to justify their expenditure by the effects on their own clientele. At best the counts should differentiate:

- Use from workstations on the libraries premises
- Use from workstations outside the libraries premises but belonging to the libraries institution or clientele
- Use from outside the institution or clientele.

Such differentiation proved difficult for non-costed services like the OPAC.

THE INDICATORS

The following list of performance indicators relies on the outcome of the project EQUINOX and the work of the ISO group for the Technical Report, as the University and Regional Library of Münster is involved in both projects. The criteria for selecting the indicators are those of the standard ISO 11620:

- informative content for management decisions
- reliability and validity of the results
- practicability in the sense of reasonable amount of effort
- comparability of the scores.

The indicators are grouped as to the following aspects:

- Market penetration
- Provision of electronic services
- Use of electronic services
- User support
- Human resources
- Costs of electronic services.

1. Market penetration
    Percentage of the population reached by electronic library services.

The library offers a range of electronic services (OPAC, databases, journals, website) to the population it is set up to serve. The indicator tries to assess whether these services are indeed used by that population.

The data for this indicator can only be gained by a user survey, whether in written or electronic form or by telephone. A sample of persons out of the population are asked, whether they have made use of the library's electronic services during a specified time period. Generally the library will add other questions in order to gain more detailed information. Such questions could be:

- what services have been used

- how often they have been used

- and from what access point (in the library, from the office, from other places).

If the indicator shows high market penetration for the electronic services, this might induce the library to stress and enlarge those services. In case of low acceptance, actions like promotion of services or user support might be taken.

2. Provision of electronic services
    Provision of electronic services can be expressed by the allocation of acquisition funds or staff resources or by offering facilities (workstations) or training lessons for users.

2.1 Percentage of expenditure on information provision spent on electronic resources
    The indicator tries to assess how far the library is committed to its electronic collection. It is especially interesting for tracing the development over years.

Expenditure on electronic resources includes subscriptions, licensing and pay-per-view charges that the library pays for its users. In case of consortia, only the library's own share in the consortial payments is counted.

The data for this indicator should be easily available in the library's acquisition statistics. A problem might occur if electronic and print versions (e.g. of journals) are acquired in a package. Usually the „normal" price will be for the print version, and only the surplus paid for the electronic access is counted as

expenditure for electronic resources. Expenditure for hardware, software and networking is not included here, though there might be expenditure directly connected with the acquisition of electronic documents.

This indicator should always be seen together with indicators about the use of electronic documents. The percentage of resources spent for electronic documents shows the library's efforts in this area. Use data show whether the resources have been spent for the material that users need.

The score will of course be influenced by the library's collection policy, especially by the subjects collected. It might therefore be informative to identify the different percentages in the subject collections, e.g. medicine or sociology.

2.2   *Percentage of library staff providing and developing electronic library services*
Offering electronic services demands a high involvement of staff, as well in user services as in background services. The indicator shows what priority the library gives to IT and web-based services.

The definition includes all library staff planning, maintaining, providing and developing IT services and web-based services. It does not include the staff in user services, e.g. reference and training services, dealing with electronic media, and the staff involved in the contents of the website. The number of library staff is calculated in FTE (full time equivalent).

Of course, certain services might be outsourced, or an external IT department might provide certain parts of the services. This should be named when publishing or comparing results.

2.3   *Number of computer workstation hours available per capita*
Libraries offer workstations for their users to provide access to their electronic services and for Internet use. The indicator assesses whether the number of workstations and the times they are accessible are adequate to users' needs.

The data needed for this indicator are:

- the number of user workstations

- the hours the workstations are accessible per year (usually the library's opening hours)

- the members of the population to be served

If library areas with workstations have different opening hours, those areas must be calculated separately and the results be cumulated. If some workstations are restricted to certain services, e. g. OPAC use, they should be calculated separately.

When interpreting this indicator, the number of computer workstations elsewhere in the institution should be taken into consideration.

A variation of this indicator would be:

> *Number of computer workstations per capita* (without considering the workstation hours)
> This variation will be less valid for assessing the degree of provision with workstations.

3.   *Use of electronic services*
     The indicators show whether services offered by the library potentially meet users' needs.

3.1 *Number of sessions on each electronic library service per capita*
     Libraries offer different electronic services: OPAC, website and the electronic collection (databases, journals, single documents). It has not proved efficient to calculate an overall indicator for this; but the indicator can give valid information about the use of each single service (e.g. a certain database).

The data needed are sometimes difficult to collect. They are:

- the sessions (accesses, log-ins) on each service by members of the population to be served during a specified time. Accesses by the population can be easily distinguished for costed services, but this might be difficult for services offered free to everybody (e. g. OPAC).

- the members of the population to be served.

The results can be used for comparing:

- the amount of use over years

- the amount of use for different services

- the use of the same service in other libraries.

It would be interesting to compare the use of a service with its costs (see 6.1).

3.2   Rejected sessions as percentage of total attempted sessions
This is an indicator that tries to assess the availability of electronic services. It measures whether there are enough licenses for users' demand. It is, therefore, only valid for costed services - but these are the important ones.

The data needed are:

- the number of rejected sessions on each costed electronic service during a specified time period (by members of the population of course, but costed services would propably be restricted to these)

- the number of all attempted sessions on the same service

Sessions rejected because of incorrect passwords or user ID's are excluded.

The indicator could be compared to a traditional indicator for the availability of print material. The question is the same: Does the user get direct access to a title in the collection the moment he wants it?

3.3   Number of documents and entries viewed for each electronic service
Users might access electronic services without finding what they are seeking for. If they deem a document or entry worth of full display, they have potentially found items of interest.

The indicator assesses the relevance of electronic services to users' interests. It should be interpreted in connection with sessions per service (see 3.1) and cost indicators (see 6.1 and 6.2).

3.4   Number of remote sessions on electronic services per capita
The indicator assesses to what degree the population makes use of the library's services from outside the library. A variation, for which the same data could be used, would be:

Remote sessions as percentage of all sessions
The data in this indicator are cumulated to obtain a global figure for all services, though of course remote use may differ for each service.

The data needed for the indicator are:

- the number of remote sessions on all electronic services by members of the population

- the members of the population.

For some services it may be difficult to differentiate between accesses from inside or outside the library, and in the case of non-costed services, to differentiate between the population and other users. Still, a high score indicates that a large number of the library's population wants to use library services from their office or home. This could influence the library's decisions on enlarging electronic services.

3.5 Percentage of information requests submitted electronically

Libraries traditionally offer information services, usually at one or several service points inside the library. User questions are apt to be delivered in direct contact, by mail, fax, or telephone. Today, a growing percentage is submitted electronically. The indicator tries to assess to what percentage users are switching to electronic means of communication.

The data for this indicator can only be collected by staff recording during a representative time period

- all information requests

- all such requests submitted electronically.

A high score may influence the library's decision to introduce online help desks or similar services.

3.6 Computer workstation use rate

The indicator assesses whether the workstations that the library offers are adequate to demand. The data are easily available, but must in most cases be counted manually at random intervals over a period of time. They are:

- the number of workstations in use at the time of investigation

- the total number of workstations provided.

Workstations out of order should be excluded. The indicator can be compared to traditional indicators like „seat occupancy": A high use rate - possibly in conjunction with queuing - indicates a need for more facilities.

4   User support
The indicators evaluate the library's training services for users.

4.1 Number of hours training on electronic library services provided per capita
Libraries offer training lessons in the use of their electronic services and the Internet. This indicator assesses the priority that the library gives to such training.

The data are:

- hours of training lessons provided per year

- the members of the population.

A training lesson is defined as a pre-planned course that could be held in-house or externally. The score should be analysed in conjunction with the use of such lessons (see 4.2).

4.2 Number of user attendance hours at electronic library training lessons per capita
The indicator assesses whether the library is successful in reaching its population by training.

The data must be collected manually at each training lesson. They are:

- number of attendance hours at all training lessons during a year

- members of the population.

A low score may indicate a need to promote the training or bad quality of the training. The indicator should be interpreted together with feedback forms of the different training lessons. A high score might indicate a need for training and lead to offering more training lessons.

5   Human resources

5.1 Number of attendance hours at IT and electronic library training lessons per staff member
The growing scope and importance of electronic library services demand qualified and engaged staff. The indicator assesses the improvement of staff skills.

The indicator counts only pre-planned lessons on IT and electronic library services. One staff member might have attended several such lessons. As lessons can differ greatly as to duration (from an hour to several days), the hours of attendance are counted.

The data needed are:

- number of attendance hours at lessons on IT and electronic services
- number of staff members (not: FTE).

The indicator does not include informal training, which is frequent, and can, therefore, only give a rough estimate of the average degree of training.

6   Costs
These indicators assess whether the provision of electronic resources is cost-effective.

6.1   Cost per session
The indicator relates the costs of an electronic resource (database, journal, single document) to its use.

The data are:

- the costs of that electronic resource during a specified time (usually a year)
- the number of sessions on that resource during the same time.

„Costs" in the sense of this indicator include acquisitions, subscription or licensing costs paid by the library. Pay-per-view costs are not included, as the costs per session are evident.

The indicator may be used for comparisons over time, between different resources or between libraries. If the indicator shows low efficiency, cancellation decisions should also consider the impact of that resource. A special database with low efficiency might be important for a small research group.

6.2   Cost per view
The indicator assesses the efficiency of an electronic resource (database, journal, single document) by comparing the number of full views to their costs.

The data needed are:

- the costs of the electronic resource during a specified time (usually a year)

- the number of views from that resource during the same time.

As viewing shows a more definite interest of the user, this indicator might give more valid indication of cost-effectiveness than „cost per session" (see 6.1).

Both indicators – „cost per session" and „cost per view" – could be validated by user surveys that ask for the frequency and the kind of use of electronic resources.

ARE THESE QUALITY INDICATORS?

Performance indicators are meant to assess the „goodness" of library services, not only the quantity. Some of the afore-named indicators do not seem to come up to this claim. They measure the amount of resources allocated to electronic services, e.g.:

- Percentage of expenditure on information provision spent on electronic resources

- Percentage of library staff providing and developing electronic library services

- Number of computer workstation hours available per capita.

Such indicators show the library's engagement in electronic services. They can be seen as indicators if the library has named as one of its main goals to offer its services - as far as possible and sensible - in electronic form. They will be valid for libraries till this goal has been reached.

Other indicators – indicators of use, of costs, of market penetration – will have long-time relevance for the digital library.

USER SURVEYS

Assessing the use of electronic services is still problematic, as shown before. User surveys can help to validate the data.

Questionnaires can be offered to users in electronic form (e.g. connected to the use of an electronic service), or as handouts, or a survey might be performed by telephone. Questionnaires should cover the following:

- demographic data (e.g. undergraduates)

- which services were used (e.g. OPAC)

- frequency of use

- location of use (inside the library or the institution, from other places)

- and what difficulties occured.

The University and Regional Library Münster conducted a telephone survey of 300 registered users[4]. The most difficult point was to reach students at their telephone addresses; the recall was high, as most persons immediately consented to be inverviewed. Some of the most interesting results were:

Services used	
OPAC	88 %
Website	60 %
Electronic journals	24 %
Databases	24 %
Document delivery	21 %

Evidently, there is much need for further propagation of some services.

No problems occured	
OPAC	75 %
Website	96 %
Electronic journals	89 %
Databases	89 %
Document delivery	90 %

The astonishing result was that users found most problems in handling the service they use most often: the OPAC. This might be influenced by the situation in Münster where there are still two online catalogues and some card catalogues.

Location of access	
the library	46 %
the university	20 %
Other	26 %
no use	8 %

Though the library is still the main point of access for the use of electronic services, a growing percentage uses electronic services from their office or home. This remote use does not concentrate on closing hours of the library, as e.g. OPAC statistics show.

WHAT SHOULD BE DONE FURTHER?

Quite a number of indicators has been developed and tested in projects. Libraries should continue to use these indicators and compare results and methods so that a general consense on validated indicators can be reached. Furthermore, it is important to contact the suppliers of information and the vendors of library systems and get them to deliver the data that we need in order to reduce time-consuming data collection.

Libraries put much effort into the development of their electronic services. We must be able to show, whether we are developing in the right direction, whether the services are accepted by our clientele, and whether they are offered in a cost-effective way.

REFERENCES

1   A former version of this paper was published in the Electronic Library Vol. 19 (2001), Nr. 5.

2   ISO 11620 (1998), Information and Documentation - Library Performance Indicators, p. 4.

3   The definitions are taking from the draft of ISO 2789.

4   Mundt, S. and Bell, E. (2000), „Daten über Daten – Telefonische Befragung von Bibliothekskunden zur Nutzung elektronischer Dienstleistungen", *Bibliothek, Forschung und Praxis*, Vol. 24 No. 3, pp. 288 – 296.