
A project to benchmark university libraries in The Netherlands

Hubert Laeven and
Anja Smit

The authors

Hubert Laeven (e-mail: h.laeven@mailbox.kun.nl), formerly Director of Libraries, is a Researcher, Consultant and Interim Manager and Anja Smit (e-mail: Anja.smit@ubn.kun.nl) is the current Director of Libraries, both at the University of Nijmegen, Nijmegen, The Netherlands.

Keywords

Benchmarking, Statistics, Performance measures, University libraries, The Netherlands

Abstract

In 1998 and 1999 a small group of Dutch university libraries started working together to develop a set of instruments for performance measurement through benchmarking. This article gives a short description of the project, the results, and the local implementation of the outcomes of the project at Nijmegen University Library.

Electronic access

The Emerald Research Register for this journal is available at
<http://www.emeraldinsight.com/researchregister>

The current issue and full text archive of this journal is available at
<http://www.emeraldinsight.com/0143-5124.htm>

Run-up to a nation-wide project

In the mid-1990s, university administrators in The Netherlands decided to found the IWI consortium (IWI = Innovatie Wetenschappelijke Informatievoorziening) (Innovation of Academic Information Services). The consortium's aim was to promote "the innovation of academic information services" by defining common lines of action, to initiate the actual realisation of this aim, and to provide the funds for all the projects involved. Not only was it the universities that participated in IWI, but other Dutch institutions with executive tasks in the field of academic information services, such as the Royal Library and the Royal Netherlands Academy of Arts & Sciences also took part. The necessary financial means to support the IWI activities and to execute the projects consisted of contributions made by the participants, and state funding. The projects' execution was assigned to (combinations of) the participating institutions through tender procedures. In IWI's first long-term plan dating from 1995, 14 project fields were defined. One of them involved the performance of a "quality audit", which was targeted at the development of the digital library and the use of corresponding national services.

In 1996 and 1997, several institutions and experts were consulted about the best way to set up such a quality audit and to develop a practical method. Additionally, an investigation took place to ascertain which institutions would be prepared to carry out a project for the development of the new method. In the end it was decided to restrict the project to university libraries and to assign the project to the libraries of the universities of Nijmegen, Rotterdam, Groningen, Twente, Leiden and Amsterdam (VU). The project group was to be supported by external experts from TNO-STB, a Dutch research consultancy working in the field of strategic issues concerning the relationship between technology and social innovation, and by the Library Research Department of the Royal Library. The project's duration would be two years, starting from January 1998. A report on the project was published by IWI (Frowein *et al.*, 2000).

The authors would like to thank Astrid Ras for translating the paper into English.



The aim of the project

IWI defined the following aims for the project:

- (1) The design and realisation of a set of instruments based on performance indicators, with which libraries can:
 - measure the effectiveness of their services in relation to the requirements of their target groups;
 - gain a better insight into the efficiency of their management; and
 - analyse the corresponding processes better and even improve them, if necessary.
- (2) The development of policy instruments for library management, consisting of performance indicators and a concise analysis which can be used to support policy decisions concerning the deployment of human and material resources.
- (3) The creation of an opportunity to periodically compare the performance of one library with that of other libraries, using the performance indicators.

Over the years, the Dutch university libraries, working closely together, had already collected statistical data. However, it proved to be difficult to compare the data collected by one library to those collected by another (Voorbij, 1993, 1994). The reason was the lack of clearly defined concepts. This problem obviously had to be examined. Furthermore, it was agreed that the measuring instrument should be used for all library processes and should not merely concern the development of the digital library, as IWI had initially advocated. The explicit intention was that, after the project had been completed, all Dutch university libraries would use the method and that the project would lead to a more or less permanent monitoring of the libraries' performance in relation to one another. Therefore the set of instruments had to be general, practical and restricted to issues for which mutual comparison was appropriate, which clearly excluded local issues. Also, it was clear from the outset that the scale of the efforts to collect data should be modest, in order to procure the participation of all the libraries concerned.

Theoretical orientation

The project group started with a three-fold theoretical orientation. Its members briefly studied ideas concerning integral quality management, the factors involved and the models that are the basis of most quality systems. Additionally, the group assessed which instruments, i.e. performance indicators, questionnaires and audits or visitations, could best be used to obtain the required information about various aspects of the business activities, and about factors influencing quality in a business environment. For both areas, i.e. for the model-oriented approach of the problem and for the differentiation into assessment instruments and measuring instruments, the project group principally chose material which is used by the Dutch Institute for Quality (Instituut Nederlandse Kwaliteit or INK) (see www.ink.nl/index_ink.htm). In addition, a brief scan was performed of the most recent literature on performance measurement in the library and information sectors (Lancaster, 1993; Baker and Lancaster, 1991; King Research Ltd, 1990; Van House *et al.*, 1990; Cotta-Schönberg and Line, 1994; Ward *et al.*, 1995; Cullen and Calvert, 1995; Poll and Te Boekhorst, 1996; ISO, 1998).

For the purpose of this article it would stray too far afield to extensively discuss the philosophy of integral quality management. Nor was this done for the purpose of the project itself, which was named "Benchmarking in university libraries". Instead, the idea was that a selection of principles from prevailing theories would suffice to support and justify the additional efforts that would result from the introduction of benchmarking in libraries.

One of the major principles of integral quality management is to constantly aim at improving one's own performance, at individual, group and organisational level. Another fundamental consideration is to safeguard the reliability of services, which is expressed in the aspiration to aim for a constant quality level that is not influenced by occasional factors. Users of libraries and information services should be able to rely at all times on equal levels of expertise and identical ranges of services at the various locations of the organisation concerned. Integral quality management is based on the fact that good intentions alone will not

provide adequate results. Organisational measures will have to be taken and exact goals will have to be defined. This implies that measurements should be taken at regular intervals in order to assess whether the goals concerned are being achieved. Subsequently, plans for improvement can be drawn up and implemented, after which new measurements should be taken to assess whether further adjustments are called for.

The third and last key principle of integral quality management is the notion that publishing the performance and quality assessments concerned will lead to the greatest possible incentive for further improvements. In the academic world this notion has been put into effect in the form of visitations, periodical and public assessments of tuition and investigations at the various universities and faculties by special, national committees of experts.

For this particular project, the quality model of the Dutch Institute for Quality (Instituut Nederlandse Kwaliteit) (INK), noted above, was chosen as the framework for the assessment and deployment of benchmarking as one of the instruments for performance measurement (see Figure 1). This model will provide the required causality between factors that are essential to a system

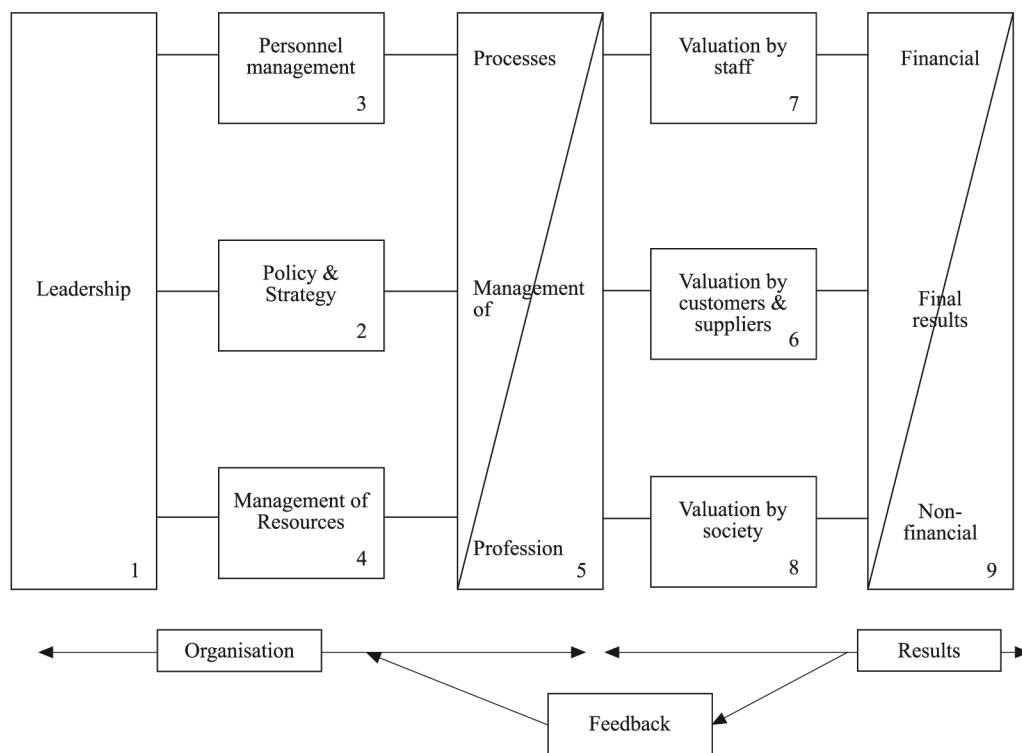
of quality management. In addition, it will also be an incentive for management and staff to take a more critical view of their own organisation, by taking into account factors such as professionalism and efficiency, thus advancing the insight into one's own failures and achievements and paving the way towards change.

The model of the Dutch Institute for Quality

Areas 7-9 of Figure 1 are the result areas. The final results, both financial and non-financial, are largely determined by the results in the other three areas. Customers, suppliers (whose role has become increasingly important because of specialisation and chain thinking) staff and society are all interested parties when it comes to an organisation's performance. The satisfaction of these parties determines the rating of the results to quite a large extent.

Areas 1-5 are so-called organisation areas. In these particular areas, the organisation should organise and structure itself in such a way that the set goals can be achieved. Process management is pivotal. The other organisation areas mainly include the means

Figure 1 The INK model



and instruments to achieve the desired results and goals.

A very important principle in quality management is constant feedback within the cycle: setting goals, executing them, measuring, executing or adjusting the goals (plan-do-check-act). Such feedback cycles can be found at three levels in the INK model.

- (1) It almost seems to be a truism to say that, in general, there is a close link between the organisation areas and result areas as such.
- (2) Furthermore, there are at least five clear-cut "bilateral" feedback connections between:
 - process management and non-financial results, such as duration, occupancy rates, productivity, waiting lists etc.;
 - policy and strategy on the one hand, and customer satisfaction on the other. Here, the added value an organisation wishes to offer its customers and the way it is achieved is of key importance;
 - leadership and valuation by society. This is mainly expressed in the organisation's social involvement and the way it gives shape to its social position;
 - management of resources and financial results, the business economics connection *par excellence*; and
 - personnel management and the valuation by the staff. The main question here is: does the organisation use the capacity of its most important resource, the staff, to the full and which view is taken by the staff on this matter?
- (3) Finally, there are of course close feedback connections between the sub-domains within most organisational areas. For instance, between a number of sub-domains of process management, such as between process identification, process control and management, process assessment and improvement, and the stimulation of innovation and process modification. The cyclic movement "plan-do-check-act" can be clearly recognised here.

Instruments for quality measurement

In the previous section it has been argued that the quality model offers a framework within which organisations can strive for quality enhancement. The model may serve as a kind of checklist for the management team that is focused on "total quality management".

Furthermore, the model will highlight the interdependence between the various aspects of business organisation and management. However, it should be taken into account that not all the items and aspects included in the model can be expressed in clear quantities. Therefore, quality is usually measured and assessed using three types of instruments:

- (1) Performance indicators for those aspects which can be expressed quantitatively in a meaningful manner.
- (2) Questionnaires for those aspects about which an opinion is required.
- (3) Audits or visitations for those aspects about which an expert assessment is thought to be required.

Usually, systems for integral quality management use all three of the above measuring instruments. For the project "Benchmarking in university libraries" not only the INK quality model was used. In order to determine which aspects could best be measured using performance indicators, the methodological guidelines prepared by the Dutch Institute for Quality (Instituut Nederlandse Kwaliteit) were also used. They can be found in Table I. It appears that performance indicators are especially applicable to the items finance and material / equipment (especially the collection) within the section management of resources, and to process management and the final results.

For periodical comparisons between the performance level of various companies and / or institutions, benchmarking according to performance indicators, is of course the most appealing method, since it is very effective and efficient. However, for a meaningful application of the method, a number of important conditions should be met. First, only similar organisations can be compared to one another. Therefore, our project was restricted to university libraries. Furthermore, performance indicators ought to be defined according to a single method. This implies that precise measuring quantities should be defined and clear instructions about the measuring methods should be given. In

Table I Types of quality measurement

Special attention area	Sub-area	Performance		Audit/ visitation
		indicators	Questionnaire	
Leadership	Involvement of the management			00000
	Promoting a quality structure			00000
	Acknowledgement and appreciation		00000	
	Actual support		00000	
Policy and strategy	Aiming for quality care in P&S			00000
	Use of information			00000
	Quality of implementation plans			00000
	Communication P&S			00000
	Assessment and improvement			00000
Personnel management	Personnel policy			00000
	Expertise			00000
	Goals and assessment			00000
	Promoting continuous improvement			
Management of resources	Funding	00000		
	Information		00000	
	Material and equipment	00000		
	Technology			00000
Process management	Process identification	00000		
	Process management and control	00000		
	Process assessment and improvement	00000		
	Promoting innovation and creativity	00000		
	Process change		00000	
Valuation by customers and suppliers	Users		00000	
	Governors			00000
	Suppliers		00000	
Valuation by staff	Right to participate		00000	
Valuation by society	Right to use		00000	
	Right to continue working		00000	
Final results	Financial results	00000		
	Non-financial results	00000		

addition, comparison between indicators should provide the management with meaningful information, so that appropriate subsequent action can be taken. Each indicator should relate to aspects of business management which can be influenced by the management, and not to circumstantial factors and the like, which either are not, or cannot be, influenced by the management in any way. In order to enhance comparability, the indicators should be set in the correct context and corrected, if necessary, for instance, according to the size of the participating organisations. Also, the number of indicators should remain manageable. Hence, they should be restricted in number and be surveyable, especially where periodical application is concerned. The restriction in number is also required in order to achieve an accurate and thorough data collection and supply. If the "shop-floor" workers feel

overburdened with collecting facts and figures, the quality of the supplied data will deteriorate. Usually, co-operating organisations that are seeking to apply a common benchmarking system restrict the number of indicators to between 15-20.

The most frequently used method is one in which a co-ordinating institution periodically collects the required data, creates tables for each indicator and calculates averages and distribution. The participants subsequently receive a survey, in which the organisation's position is set against that of the colleagues and against its own previous position, per individual indicator. A brief analysis is also added. The organisation can then compare the indicators with the set goals and take corrective action if required. The nature of the branch of industry concerned determines the extent to which anonymity is guarded when the data is presented.

When participating in a benchmarking system, success largely depends on five factors:

- (1) The extent to which the method will support the management in realising quality management within the organisation itself.
- (2) The extent to which the method will offer added value when compared to internal management reports.
- (3) The possibility to identify opportunities and bottlenecks as objectively as possible.
- (4) The possibility to test the organisation's own course against the performance of others and against the organisation's own former performance.
- (5) The extent to which the method will contribute to a greater accountability to the outside world with regard to the organisation's goals and performance.

Exploration of the literature

The following conclusions can be drawn from literature that was explored for the purpose of this project.

The composition of a set of performance indicators may proceed along two lines. The indicators are either derived from a theoretical model or selected for pragmatic reasons (inductive versus deductive). The two approaches are by no means mutually exclusive. A first selection on pragmatic grounds may be followed by a test against a theoretical model. This then becomes a control mechanism to assess whether all aspects and angles are sufficiently represented in the selection.

Moreover, most manuals are not restricted to listing indicators that are based on statistics and data collected and provided by the library. Research on library users is also commonly recommended as an evaluation method. Entirely in accordance with the current climate of focusing on the user, direct user information is regarded as being essential to a good and comprehensive assessment of quality.

Performance measurement mainly focuses on those aspects that are best suited to be measured and which can be expressed in exact figures. This is confirmed by the overview of measuring methods presented above, which has been obtained from the Dutch Institute for Quality (Instituut

Nederlandse Kwaliteit). The development of indicators for electronic libraries was still in its infancy when this project was executed. In the meantime, considerable progress has been made in this field. The project E-metrics by the Association of Research Libraries (see www.arl.org/stats/newmeas/emetrics/) seems to play a major part here, while the International Standard Organisation has also made a valuable contribution, which is proven by ISO 2789. Recently, the first results of the project have been published in a series of five documents (Association of Research Libraries, 2002). The term "benchmarking" was not then used in the literature studied. In fact, performance measurement was discussed without making a real attempt to compare data. In recent years, this attitude has been changing rapidly, which is clearly demonstrated by today's international interest in benchmarking for libraries.

The stages of the project and outcomes

The project, in which six of the 13 Dutch university libraries participated, was sub-divided in six stages:

- (1) A brief study of literature.
- (2) Composition of the first version of a set of performance indicators (36 altogether).
- (3) A test of this set of indicators by the six libraries.
- (4) Revision of the set of performance indicators according to experiences during the test. This resulted in a set of 26 performance indicators and two questionnaires.
- (5) Test of the revised instruments.
- (6) Composition of the final set of 24 performance indicators and two questionnaires.

The first stage was to compose a list of performance indicators based on the examination of the literature. The set was regularly tested against the INK model in order to check for sufficient coverage and distribution among the various categories. Intense discussions led to a narrowing down of the original set to 36 items. The main criterion was the practical significance to the management. After all, libraries also collect data for local purposes. In the eyes of the participants, the (comparable) data to be collected for the purpose of the project should

possess added value over local statistics. From the outset, the idea was that data should be comparatively easy to collect. In two test rounds, combined with evaluation and discussion rounds, the initial list of 36 indicators was reduced to 24 items. Furthermore, the participants became increasingly convinced that the set of instruments would be incomplete without questionnaires to complement the performance indicators. Hence, seven questionnaires were developed, two of which would be part of the standard set of instruments. The others were marked as being optional, to be used by libraries according to their own requirements. The indicators were sub-divided into four groups:

- A. Financial and human resources (primary input).
- B. The range of facilities for the user (secondary input).
- C. Efficiency of internal processes (throughput).
- D. The use of the various facilities (output).

The performance indicators

An overview of the final set of 24 performance indicators is listed below.

Library resources

- A1 Library expenses/university expenses.
- A2 Library's own revenues and expenditure.
- A3 Expenses for collection development/library expenses.
- A4 Expenses for subscriptions to printed publications/expenses for printed monographs and subscriptions in the categories humanities, social sciences, natural sciences and total.
- A5 Expenses for electronic sources/expenses for printed and electronic sources, in the same categories as under A4.
- A6 Distribution of FTE library staff between
 1. Document acquisition and processing,
 2. Services to the public,
 3. Management and support.
- A7 Expenses for education and training per FTE of library personnel.

Library facilities

- B1 Expenses for collection development per academic member of staff, expressed in persons and FTEs, and categorised in printed and electronic monographs,

printed and electronic subscriptions, total expenses for collection development (including old prints and the like).

- B2 Number of book purchases per academic member of staff, expressed in persons and FTEs and categorised in humanities, social sciences, natural sciences and total.
- B3 Number of current (printed and electronic) subscriptions per academic member of staff, expressed in persons and FTEs, in the same categories as under B2.
- B4 The number of hours during which the central library is open weekly.

Process efficiency in the library

- C1 Number of processed book titles/ FTEs involved in book processing.
- C2 Duration of receipt - availability of books, expressed in the number of days on which 50, 80, 90 per cent is available respectively.
- C3 Duration of honouring requests from the central library stacks, expressed in the number of minutes in which 50, 80, 90 per cent are available respectively, as well as the percentage which becomes available within 24 hours.
- C4 Average duration ILL books.
- C5 Average duration ILL items.

Library use

- D1 Loans in the past year/acquisitions over the past five years.
- D2 Loans per borrower category, i.e. average number of loans per student of the home university, average number of loans per home university member of staff (in persons and FTEs), lending percentage for external borrowers.
- D3 Number of book loans to home users/ number of book loans to home users + through ILL.
- D4 ILL book requests received by the library/ ILL book requests made by the library.
- D5 ILL article requests received by the library/ILL article requests made by the library.
- D6 Honouring percentage concerning the ILL book requests received by the library.
- D7 Honouring percentage concerning the ILL article requests received by the library.
- D8 Number of instructions, training sessions, demonstrations.

Discussion

Of seven indicators in the category "Library resources", two were concerned with the means of acquisition, four were concerned with the input of resources and one the distribution of the FTEs among the departments "Processing", "Public services" and "Management and support". With regard to the resources for collection development a distinction is made, for obvious reasons, between printed material versus electronic material on the one hand, and the main disciplines humanities, social sciences and natural sciences on the other. The indicators in this category are not so much concerned with performance as with the recognition of patterns and trends, which is extremely relevant for the management. For practical reasons, only library and university expenditure have been included, since determining the respective budgets yielded huge problems and differences in interpretation. The indicator "resources per capita" has been omitted, for it seemed too broad. And as for the indicator "expenses for collection development/total library expenses", the various regimes regarding housing costs proved to stand in the way of any useful comparison.

The category "Library facilities" consists of four indicators, three of which concern collection development. This makes the category somewhat unbalanced. Several other indicators, however, were eliminated on the basis of the tests. For instance, the availability of PCs for end-users seemed to depend too heavily on local circumstances: the distribution across buildings on campus versus the concentration in the library, the availability of facilities for teleworking versus facilities on campus etc. Additionally, the count of seats provided was equally dismissed, as measuring the occupancy rate was considered to be too labour-intensive. Another four indicators underwent the same fate, viz. the number of students per available seat, the number of electronic periodicals as opposed to the total number of periodicals (the expenses for electronic sources were considered to be sufficient), the number of books on public display as opposed to the number of acquisitions in the past year, and the percentage of periodical volumes on public display.

Only one indicator from the category "Process efficiency in the library" is aimed at

productivity. This is the number of processed book titles against the total number of FTEs in the sector book processing (the total of acquisition, cataloguing, allocation of shelf marks and subject cataloguing). During the project one of the more restricted indicators, viz. the productivity of cataloguing, was cancelled, as the responsibilities of the various cataloguing departments in the participating libraries proved to be rather divergent.

Initially, the number of acquisitions per capita had also been taken into account, in addition to the number of processed acquisitions. This indicator was struck off in order to reduce the data collecting workload. It also proved to be rather difficult to reach uniformity concerning the method to collect data.

The remaining indicators in this category concern processing times. These indicators were considered to be very important because they affect users, and also because of the fact that high productivity is not necessarily related to satisfactory processing times.

In the category "Library use", three indicators concern loans, four concern inter-library lending, and one concerns information to end-users. The first indicator, the number of loans from the past year divided by the number of acquisitions from the past five years, is meant to produce an overall insight into the effectiveness of collection development. Two alternatives were struck off the list. One was the average lending frequency per volume from the collection: the outcome is usually relatively low, as university libraries have quite a lot of older material in their collections. The other alternative, the average lending frequency of recent material, albeit a much more interesting indicator, is difficult to measure. Finally, an attempt was made to gain insight into the average number of loans of recent acquisitions and the number of recent acquisitions that have not been borrowed. However, the sizeable sample surveys that would be required and carried out manually would produce too great a workload.

The second lending indicator, the average number of loans per borrower category, was initially complemented by asking about the number of active borrowers, which was intended to give an insight into the distribution of the loans. For technical reasons (usually, lending systems are unable to avoid double counting of borrowers who

are linked to loans) this indicator had to be eliminated.

The third indicator is intended to provide an insight into the extent to which a library's own collection meets user requirements. Initially, this indicator had a counterpart for periodicals. However, for various reasons it proved to be too difficult to apply.

The remaining indicators concerning inter-library exchange mainly represent the role that the various libraries are playing nationally in the field of document supply.

Additionally, an attempt has been made to measure the use of electronic full-text files in various ways, in order to monitor trends over the years. However, not a single library proved to be capable of producing the required data adequately. Therefore, these data have been eliminated (for the time being). The questionnaire "Reach of the library", however, may be an adequate substitute in this respect.

The only non-collection related indicator in this category is the number of instruction sessions provided by the library. For reasons of simplicity, sessions were preferred to hours. A "per capita" indicator did not seem to be useful here. The category's composition is therefore rather unbalanced, but this is somewhat compensated by the addition of data from the questionnaire "Reach of the library".

The questionnaires

As has been demonstrated, not all aspects of business management can be assessed using performance indicators. For matters on which the opinion of the (potential) user is essential, questionnaires are the better instrument. For this project, two questionnaires were chosen as part of the standard set of instruments for benchmarking in the Dutch university libraries.

The aim of the first questionnaire was to measure the library's reach: who uses which facility and for what purpose; who does not and why. The second questionnaire was intended to map customer satisfaction concerning the facilities and services provided. The questionnaires can be completed within five minutes and libraries are allowed to add their own questions, which may be important to the local situation. It is recommended that the questionnaires be

completed in a sample survey in which at least 500 to 700 respondents are participating. In this group, students and university staff should be proportionally represented. The results will have to be assessed in connection with the indicators.

The model of the Dutch Institute for Quality (Instituut Nederlandse Kwaliteit) recommends the use of a larger number of questionnaires in order to map an organisation's "critical success factors" and their mutual relations more comprehensively. Except for the two aspects mentioned above, such questionnaires are also recommended for the factors of leadership, management information as part of the management of resources, introduction of process change, valuation by suppliers, valuation by staff, valuation by society. Therefore, the project organisation decided to develop questionnaires for these elements as well, although their application was marked as being optional. The most important reason for doing so was the fact that they mainly concern quality and are more suitable for internal purposes than for mutual comparison, while the principal two questionnaires contain essential additions to the more quantitative parts of the method.

The set of instruments and their application

The complete set of instruments available to the libraries after the completion of the project consisted of the following elements.

- (1) A manual on how to supply data. It contains final directions about the method for data collection, for all "measuring quantities". The measured quantities have to be fed into an Excel spreadsheet called "Measuring results". Additionally, a (Word) form called Comment measuring results is available, in which a comment about the measured results can be given (exceptional circumstances, problems during the data collection, suggestions for improvement etc.).
- (2) An overview of the measuring quantities and performance indicators. After having fed the measured quantities into the Excel form Measuring results, the performance indicators will be automatically presented.

- (3) Guidelines for interpreting the indicators. This concerns the significance of the indicator concerned. The main issue is which policy decisions could be considered on grounds of a low or a high score. Furthermore, an attempt is made to provide an insight into the “context and determinants” of the various indicators. First and foremost, this allows for a better assessment of the library’s own scores. Also, the interpretation of the scores may call for further investigation, in case one of the participants has excellent scores. It is entirely in the spirit of the benchmarking philosophy that the “best practices” which may result from it will yield positive effects for the group of participants as a whole.
- (4) Questionnaires. They will provide a better insight into those aspects of business management which concern quality. For the purpose of the project, eight questionnaires have been designed, two of which are part of the standard set of instruments and the other six labelled as being optional.
- (5) The form “Results benchmarking”. The co-ordinating institution incorporates the assembled data in a survey, which is then returned digitally to the participants. In the meantime, the Dutch libraries have passed on the assembling and processing of the data to the Royal Library. The total survey should at least contain:
- value per indicator for the library in question over the past three periods;
 - value of the indicator for the other libraries in the period concerned; and
 - average value of the indicator over the past three periods.

The results will subsequently be briefly commented upon by an expert from the Royal Library. This person will also be responsible for any adjustments to the set of instruments in the next round.

At the end of the project, the UKB, the co-operation of university libraries, the Royal Library and the Library of the Royal Netherlands Academy of Arts & Sciences, decided that, from the year 2000 onwards, all university libraries shall annually send their data for national benchmarking to the co-ordinating centre. Everyone was aware of the fact that the set of instruments developed was rather limited and that there were still many desiderata. This

was also due to the fact that a pragmatic approach had been explicitly opted for: the exercises should not be too time-consuming and the data collection had to be in close keeping with everyday practice in the test libraries. Especially with regard to the insight into the development of the electronic library, this still proved to be too much of an obstacle.

The first joint round was difficult, mainly because data on 1999 were required. Additionally, the fact that half of the libraries had not participated in the project slowed things down. These libraries had not yet had the opportunity to adapt their internal statistics and data collection to this new type of co-operation.

Meanwhile, a national study group from UKB has been established for further development of the method. This study group has started with a sequel to the project. It had become evident that the role and performance of electronic services in the entire process deserve much more attention. Furthermore, an attempt will be made to eliminate any imbalances and omissions from the system.

Local implementation in Nijmegen University Library

Nijmegen University Library[1] was one of the libraries participating in the benchmarking project. This proved to be an advantage in many ways, one of which undoubtedly was that we were able to inform our staff at an early stage and create a climate of assessment during the project. In 1998, we formed a study group consisting of members of different departments (back and front offices). We felt it was important to bring together expertise concerning day-to-day routine, processing data and insight in the library’s main goals. This local project group discussed performance indicators, proposed changes and instructed colleagues on the collection of data needed for the tests. As the deputy librarian was both a member of the national project group and the local project leader, she was able to monitor progress on two fronts.

Of course, the definitions for “high performance” and “good quality” were discussed in both groups, and it was very difficult to reach consensus. Is performance good or poor, when your circulation figures are lower than those of other libraries, while at the same time 50 per cent instead of 75 per

cent of your collection is housed in closed stacks? Situations may vary, goals and missions of libraries may vary as well. And although university libraries seem to be comparable, when you take a close look, an amazing variety comes to light. It soon became apparent that every group or department involved in the benchmarking project had to go through the same process of discussing the concept of "quality" before it could move on and get to work. Eventually, discussions always ended in the same way: we can only conclude that any choice for a definition of high performance is arbitrary to a certain extent. It is difficult to measure everything to the degree you would like to measure it, and not all measurements are of strategic importance. Nevertheless, we strongly recommend taking ample time for this process, as it is an important step in the acceptance of a "climate of assessment".

In addition, the local benchmarking project group screened the entire existing statistical practices of the library. An effort was made to delete outdated measures, to propose new ones, to replace some measures by the ones needed for the benchmarking project, and to develop a blueprint for the collection of data. Also, a new reporting tool was implemented. The most important issues during the entire process were the incorporation of the data needed for the benchmarking project into our local statistical system, and the optimal assurance of continuity in the comparability between the new and the old system. In 2001 we implemented the new statistics. All data to be collected, including the data collecting methods, were described in a manual. All departments used a simple Excel-sheet to supply their data, as shown in Table II.

During the entire process, which took more than two years altogether, our main motto was: be practical! If it is not possible to do exactly what you would like to do, do what is possible. If your library system cannot supply the management information you need, let people collect the data you need manually.

The implementation of new statistics in itself also takes time: automated systems can sometimes only deliver the required information after new input methods have been implemented, which often can only be introduced at the beginning of a new cycle. We found that two years preparation time is a realistic period for the introduction of new measures.

Actions such as updating the manual, accentuating definitions and fine-tuning data collection methods or measures are foreseen annually. The action of processing and integrating the data is centralised: a management assistant is responsible for this part of the process. The data set will include input data for benchmarking, but also measures to be published in reports, such as the annual report. The first-time gathering of new data was very time-consuming; most of all because the new methods had to be tested first and some improvements had subsequently to be made. Two years later, the process took far less time and energy than the first time. By now, it has become part of an annual routine and it appears to be saving time by comparison with the former organisation of our statistics.

Improving performance

The process of revising our statistics raised awareness among our staff that measuring performance is important. This has been stimulated by the presentation and publication of the benchmarking results, in combination with annual discussions about our own local "trends and developments". All departments (back and front offices) have to prepare a presentation regarding particular developments in their domain, using the statistics and performance indicators they have collected themselves. This way, data collection will acquire greater significance for the people who have to supply the data.

The book processing departments have been asked to constantly monitor the number of days needed for new books to become available to the users. The method by which we keep track of our performance is a fine example of practicability: we use red paper strips, which are put (at random) into 50 books by the acquisitions department every month. Each department handling these items will make a note of the date and hour of receipt or transfer (to the next department) on these strips. By analysing these data, we know where bottlenecks are occurring. Problems, if any, will be discussed with the departments involved. We found that the simple fact of making performance visible greatly helped to speed up processing times.

Of course, some members of staff initially treated these clearly marked books differently

Table II Monthly report of a branch library

Monthly report branch library	1	2	3	4	Total
Stand-alone CD-ROMS					
Items available					5
New items	1	0			1
Items removed	0	0			0
Total					6
Books					
Items available					41,194
New items	438	116			554
Items removed	5	17			22
Total	443	133			41,726
Periodicals					
Items available					13,980
New items	90	464			554
Items removed	312	407			719
Total	402	871			13,815
Average number of visitors in the evenings and Saturdays	9	8			8.5
Instructions					
Number of instructions	0	13			13
Number of students	0	126			126
Number of hours	0	182			182
Information desk					
At the desk					
Number of questions	160	169			329
Number of clients	158	166			324
Through telephone					
Number of questions	4	4			8
Number of clients	4	4			8
Through e-mail					
Number of questions	7	5			12
Number of clients	7	5			12
Number of sessions online hosts	0	0			0

from the other books, so that the data were not entirely representative. But after a while "social control" normalised the situation. It would have been even better if our library system had been able to produce this type of information for all processed items. However, it was soon recognised that this was impossible, so we chose the manual method, taking the above-mentioned risk into account. But even if performance based on the marked items is not fully representative of our overall performance in this area, we may at least expect a certain consistency in the handling of these books over the years. We therefore compare the outcome annually, considering it to be reliable information, and act on it, for example by redefining our goals for the next year. Moreover, a selection of the results is also being published in our quarterly and annual reports, and subsequently presented to our users in the *Library News*.

Cullen and Calvert (1995) demonstrated that library staff are not very good at predicting what users (and students in particular) regard as being characteristics of a highly effective library. In 2002, in the framework of a large scale user survey, we therefore decided to ask our users to give their opinion on some services which are also measured through performance indicators, e.g. the processing time for honouring requests from the closed stacks in the Central Library. The results of the user survey may redirect our efforts or affect our system of collecting data in one way or another.

The significance of benchmarking in everyday practice

The characteristics of our library, compared to the other university libraries in The Netherlands, have not changed dramatically

since the introduction of benchmarking. The overall picture has more or less stayed the same. Nevertheless, changes that might not greatly influence the general picture may certainly be of local interest. Therefore we would like to recommend the gathering of all data, not only those needed for benchmarking purposes, on an annual basis. As we have demonstrated, there are also practical reasons for this: an annual routine will reduce the amount of time needed for data collection and it has a positive influence on the behaviour and critical sense of the staff.

In conclusion, we would like to briefly discuss our experiences with, and ideas on, the practical use of benchmarking to date. Obviously, this will not be a complete survey on the advantages of benchmarking as an instrument for quality management, but rather an impression of a series of positive experiences over the relatively short period that has passed since the introduction of the benchmarking phenomenon in our library.

It is clear to everyone within our institution that the funding and performance of the library have improved substantially over the past few years. It was not the direct result of benchmarking, but benchmarking indisputably played an important supporting role in achieving these results. As a result of benchmarking, we have been able to make a more convincing case that the funding of the Nijmegen library, especially in the area of collection development, left much to be desired. This has resulted in a substantial rise in the acquisition budget starting from 2003. The additional funds should especially benefit an expansion of the electronic collection. Also, it has become evident that the literature supply for some faculties is lagging further behind than for other faculties, which should inevitably lead to a "rebalancing" of resources.

Similar effects can be reported for the library's opening hours, the availability of personal computers with Internet access and the range of services in general. The wishes of local users are either justified, strengthened or weakened, to a greater or lesser extent by comparison with other libraries.

The publication of the benchmarking results provides primary users with a better and more realistic image of their library and of the quality and efficiency of the services provided. They tend to have more realistic expectations and are more inclined to modify

their criticisms and requirements. It will also be easier to convince them that scarce means ask for firm choices when it comes to the range of services and their priorities. This is how benchmarking can contribute positively to the debate with users from all levels.

Furthermore, the national consultations with fellow libraries are also benefiting from the results of benchmarking, since common priorities are recognised sooner. This will, in its turn, make it easier to take joint action to position the libraries in the social field of force, to acquire project funds or to urge the authorities and their corresponding institutions to support or fund certain developments. Additionally, in negotiations between co-operatives/consortia and service industries/suppliers, some data may play a valuable role. Benchmarking will also make it possible, to a certain extent, to compare Dutch libraries to their foreign counterparts.

Perhaps the greatest effect of benchmarking can be felt within the library organisation itself. Participating in benchmarking has promoted, as has been noted earlier, a climate of thinking about quality, of assessing one's own performance and taking responsibility for it, thus improving customer relations and promoting self-criticism. This provides a sound base for a (lasting) appetite for change and for a continuous drive to enhance quality. Such an attitude will not remain restricted to the library management or to primary tasks, but - if the method is applied correctly - will also have a profound effect on the lower strata in the library hierarchy and on secondary tasks. For instance, if certain data have to be gathered in a painstaking manual process because the library system is not up to the task, this will lead to demands to improve the library system concerned. And it is self-evident that national comparison will accurately reveal any leeway in the supply of digital information. In a similar manner, the system will also generate extremely useful input and support for the strategic policy of the library, as has already been demonstrated at the beginning of this final section.

We conclude by stating that the development of a benchmarking system is no small undertaking and that the set of instruments used by the Dutch libraries is still far from perfect, but that, in our view, the value of benchmarking as a proven tool to achieve quality management should be rated very highly indeed.

Note

- 1 Nijmegen University is a medium-sized university with 15,000 students, eight faculties and a teaching hospital, five renowned research institutes, a number of affiliated institutes, of which some specialise in the study of Catholic life, history and heritage in The Netherlands. It was founded in 1923. The library consists of a central library and eight faculty libraries. It employs about 100 FTE staff, has a budget of some 10 million Euros, a collection of about 2 million volumes and about 7,500 subscriptions.

References

- Association of Research Libraries (2002), *Measures for Electronic Resources (E-metrics)*, Parts 1-5, Association of Research Libraries, Washington, DC.
- Baker, S.L. and Lancaster, F.W. (1991), *The Measurement and Evaluation of Library Services*, 2nd ed., Information Resources Press, Arlington, VA.
- Cotta-Schönberg, M. and Line, M.B. (1994), "Evaluation of academic libraries: with special reference to the Copenhagen Business School Library", *Journal of Librarianship and Information Management*, Vol. 26 No. 2, pp. 55-69.
- Cullen, R.J. and Calvert, P.J. (1995), "Stakeholders perceptions of university library effectiveness", *Journal of Academic Librarianship*, Vol. 21 No. 6, pp. 438-48.
- Frowein, J.C., Laeven, A.H. and Voorbij, H.J. (2000), *Benchmarking van Wetenschappelijke Bibliotheken*, Stichting SURF/IWI and UKB, Utrecht.
- ISO (1998), *ISO 11620: Information and Documentation – Library Performance Indicators*, International Standard Organisation, Geneva.
- King Research Ltd (1990), *Keys to Success: Performance Indicators for Public Libraries: A Manual of Performance Measures and Indicators*, HMSO, London.
- Lancaster, F.W. (1993), *If You Want to Evaluate Your Library*, 2nd ed., Library Association, London.
- Poll, R. and Te Boekhorst, P. (1996), *Measuring Quality: International Guidelines for Performance Measurement in Academic Libraries*, IFLA Publications 76, Saur, Munich.
- Van House, N., Weil, B.T. and McClure, C.R. (1990), *Measuring Academic Library Performance: A Practical Approach*, American Library Association, Chicago, IL.
- Voorbij, H.J. (1993), *Statistische Gegevens UKB-bibliotheken over de Jaren 1990 en 1991*, SI, s.n.
- Voorbij, H.J. (1994), "UKB-bibliotheken in cijfers", *Open. Vaktijdschrift voor Bibliothecarissen, Literatuuronderzoekers, Bedrijfsarchivarissen en Documentalisten*, Vol. 26 No. 11, pp. 350-5.
- Ward, S., Sumsion, J., Fuegi, D. and Bloor, I. (1995), *Library Performance Indicators and Library Management Tools*, European Commission, DG XIII, E-3, Brussels.

Further reading

- Voorbij, H.J. (2000), "Een benchmarking instrument voor de UB: kwaliteitsverbetering door onderlinge vergelijking", *Informatie Professional*, Vol. 4 No. 12, pp. 14-17.