

# Das Wagnis des Benchmarkings

## BIX im Vergleich

von Roswitha Poll



Am 17. Februar 2006 trafen sich Vertreter der IFLA-Sektion Statistik und Evaluierung und des UNESCO-Instituts für Statistik in Montreal, um über eine neue global anwendbare Bibliothekstatistik zu diskutieren. Natürlich wurden zunächst die Anforderungen an eine solche Statistik zusammengetragen. Die IFLA wünscht sich „robuste“ Daten, mit denen sie z.B. im Rahmen von WSIS (World Summit on the Information Society) die Bedeutung von Bibliotheken nachweisen kann. Das UNESCO-Institut möchte die Rolle der Bibliotheken für Kultur, Bildung und Lernen vor allem in den Entwicklungsländern ermitteln. Die neue Statistik soll daher

- die traditionelle wie die elektronische Informationswelt berücksichtigen,
- die Wirkung von Bibliotheken auf Kultur, Bildung und Lernen nachweisen,
- Vergleiche zwischen Ländern ermöglichen,
- und trotzdem kurz und übersichtlich und auch in Entwicklungsländern leicht erhebbar sein.

„Den lieb ich, der Unmögliches begehrt.“ (Goethe, Faust)

Sind die Anforderungen an ein gemeinsames Benchmarking von Bibliotheken einfacher?

---

**„Den lieb ich, der Unmögliches begehrt.“**  
(Goethe, Faust)

---

### Leistungsmessung: Die Indikatoren

Leistungsindikatoren und Nutzerbefragungen sollen die Qualität der Dienste in Bibliotheken ermitteln, sollen zei-

gen, ob eine Bibliothek „gut“ ist. Gute Bibliotheksleistung zeigt sich in Nutzerorientierung, Aktualität, Schnelligkeit, Verlässlichkeit, leichter Zugänglichkeit, kosteneffektiven Prozessen und adäquaten Hilfsdiensten wie Beratung und Schulung.

Leistungsindikatoren stellen meist mehrere statistische Daten in Beziehung, z.B. Bibliotheksbesuche pro Mitglied der Zielgruppe. Solche Indikatoren sollen folgenden Kriterien entsprechen.<sup>1</sup>

- Aussagekraft (für die Identifikation von Problemen und Managemententscheidungen)
- Zuverlässigkeit (bei wiederholter Anwendung)
- Gültigkeit (statistische Validität)
- Angemessenheit (für den Zweck)
- Praktikabilität (geringer Erfassungsaufwand).

Das sind bereits recht hohe Anforderungen. Die Schwierigkeit erhöht sich natürlich, wenn ein ganzes Set von Indikatoren von einer bestimmten Gruppe von Bibliotheken gemeinsam verwendet und die Ergebnisse auch für Vergleiche genutzt werden sollen.

### Die Suche nach den Kennzahlen

Wenn eine einzelne Bibliothek ihre Leistungen wiederholt mit den gleichen Kennzahlen untersucht, dann kann sie dadurch Probleme feststellen, Entwicklungen beobachten und den Erfolg von Maßnahmen prüfen. In vielen Fällen werden aber die Ergebnisse der Erhebungen nur schwer einzuordnen sein.

Dass man mit nur 85 % korrekt aufgestellter Bücher nicht gerade gut dasteht, ist noch leicht einzuschätzen. Aber sind zehn Entleihungen pro Mitglied der Zielgruppe nun viel oder wenig? Liegt ein Mitarbeiter mit 2000 bearbeiteten Medien im Jahr hoch oder

<sup>1</sup> ISO 11620, 1998: Information and documentation – Library performance indicators

eher niedrig? Erst der Vergleich mit anderen Bibliotheken gibt die Erläuterung zu den eigenen Daten und ermöglicht eventuell sogar Anhaltzahlen für Planungen und Ressourceneinsatz. Nicht nur die Bibliotheken, auch die Unterhaltsträger möchten Statistiken und Kennzahlen von Bibliotheken gern im Zusammenhang und Vergleich sehen. Daher kam es häufig auf Anregung übergeordneter Institutionen dazu, dass sich Bibliotheken – oder Bibliotheken mit Vertretern der Unterhaltsträger – zusammensetzten, um eine verbindliche Liste sinnvoller Qualitätskennzahlen zu finden.

Beim Start eines Projekts scheint alles ganz leicht. Ian Winkworth berichtet über die Entstehung der SCONUL-Indikatoren: „There was initial expectation and enthusiasm that we could rapidly agree on a small set of indicators that would support and fine-tune subjective judgements about the quality of libraries. By the early 1990s, after several years’ efforts, the Advisory Committee faced the common truth that the exercise would be more technically complex and politically difficult than expected.“<sup>2</sup> Immerhin dauerte es sechs Jahre bis eine Liste von sechs Indikatoren Zustimmung fand – ein Jahr pro Kennzahl!

Die Literatur zeigt, dass bei Beginn eines solchen Projekts stets etwa die gleichen Anforderungen formuliert wurden. Die Auswahlliste von Indikatoren sollte

- das gesamte Spektrum der Bibliotheksleistungen abbilden,
- elektronische wie traditionelle Dienste berücksichtigen,
- Bedeutung und Wirkung von Bibliotheken zeigen,
- Vergleichbarkeit ermöglichen (aber nur mit Bibliotheken ähnlicher Aufgabe und Struktur!),
- keine Bibliothek benachteiligen,
- Besonderheiten berücksichtigen (jede Bibliothek ist einmalig!),
- Ergebnisse liefern, die für Unterhaltsträger, Öffentlichkeit und selbst Politiker leicht verständlich sind
- und trotzdem aus wenigen Daten bestehen, die weitgehend aus der üblichen Statistik ableitbar sind.

Erstaunlich ist, dass dennoch eine Reihe solcher Listen entstanden ist, die zum Teil sogar regelmäßig genutzt werden.

### BIX im Vergleich

Der BIX ist ein Instrument zur gemeinsamen Qualitätsbestimmung von Bibliotheken. Seine Besonderheiten sind, dass

- die Durchführung nach anfänglicher Förderung durch eine Stiftung ganz in der Hand eines Bibliotheksverbandes liegt,
- die verwendeten Indikatoren nach Aspekten (Zieldimensionen) zusammengefasst sind, die beim BIX-WB der Balanced Scorecard entsprechen,
- die Teilnahme freiwillig ist,
- die Indikatorenliste keine beliebige Auswahlliste ist, sondern verbindlich für die Teilnehmer,
- die Erhebung jährlich durchgeführt wird,
- ein Ranking nach Bibliotheksgruppen erfolgt.

Vor allem durch die Struktur der Balanced Scorecard unterscheidet sich der BIX wesentlich von anderen Projekten. Die Balanced Scorecard wird zwar in einer Reihe von Einzelbibliotheken angewendet<sup>3</sup>, nicht aber für gemeinsame Indikatorenlisten.

Für den Vergleich mit dem BIX wurden nur Projekte herangezogen, die mit einer verbindlichen Liste von Kennzahlen arbeiten und tatsächlich Indikatoren, nicht nur statistische Einzeldaten verwenden. Außerdem sollten es keine Projekte sein, bei denen die Kennzahlen ohne Bibliotheksbeteiligung

<sup>2</sup> Winkworth, Ian: Performance measurement of U.K. university libraries. In: ARL Bimonthly Report 207 (1999) [www.arl.org/newsltr/207/ukperfmeas.html](http://www.arl.org/newsltr/207/ukperfmeas.html)

**easyCheck**  
library technologies

**RFID: Diese Technik macht es möglich.**

mehr Service  
mehr Sicherheit  
mehr Komfort

- 1 Automatisierte Rückgabe
- 2 Mehr Zeit für Beratung durch effizienten Personaleinsatz
- 3 Selbstverbuchung
- 4 Einarbeitung und Verbuchung am Personalarbeitsplatz
- 5 Datenerfassung mit mobilem Handhelder
- 6 Intelligenter Durchgangsleser

- RFID-Technologie
- Selbstverbuchung
- Mediensicherung
- Projektmanagement
- Beratung und Service

[www.easycheck.org](http://www.easycheck.org)

EasyCheck GmbH & Co. KG  
Steinbeisstraße 12  
73037 Göppingen

Telefon 07161/808600-0  
Telefax 07161/808600-22  
E-Mail [mail@easycheck.org](mailto:mail@easycheck.org)

vom Unterhaltsträger bestimmt wurden. Abgebrochene Projekte (z.B. bei den finnischen Universitätsbibliotheken) oder Vorschlagslisten von Leistungsindikatoren zur beliebigen Verwendung (z.B. Norwegen) wurden nicht berücksichtigt. Verglichen wird mit:

#### CASL (Council of Australian State Libraries)<sup>4</sup>

- Öffentliche Bibliotheken, 1998 ff.
- Die Daten sind nur kumuliert veröffentlicht, nicht für die einzelne Bibliothek
- Die Statistik wird in den Territorien erhoben
- Ranking nur zwischen den kumulierten Werten der Territorien
- Acht Indikatoren

#### Schwedisches Qualitätshandbuch<sup>5</sup>

- Swedish Library Association
- Alle Bibliothekssparten
- Drei-Jahres-Projekt 2002-2004
- fast 60 teilnehmende Bibliotheken
- Die Liste ist zwar nicht verbindlich, wird aber weitgehend von den Teilnehmern genutzt
- Kein Ranking, die Daten der einzelnen Bibliotheken waren nur den Teilnehmern zugänglich
- Zwölf Indikatoren

#### HELMS (UK Higher Education Library Management Statistics)<sup>6</sup>

- SCONUL (Society of College, National and University Librarians)
- 174 Bibliotheken
- 1997/98 ff.
- Kein Ranking, aber Veröffentlichung aller Daten der einzelnen Bibliotheken
- Sechs Indikatoren und Hintergrunddaten

#### Benchmarking-Projekt Niederlande<sup>7</sup>

- UKB (Verband der niederländischen Universitätsbibliotheken, der Königlichen Bibliothek und der Bibliothek des NIWI)
- 13 Universitätsbibliotheken
- 1999 ff.
- „Zieldimensionen“: Ressourcen, Produkte und Dienste, Effizienz der Prozesse, Nutzung

- Kein Ranking, die Daten der einzelnen Bibliotheken sind nur den Teilnehmern zugänglich
- 24 Indikatoren und zwei Nutzerbefragungen

### Indikatoren der Projekte

Die in den Projekten sowie in BIX-ÖB und BIX-WB genutzten Kennzahlen wurden den Perspektiven der Balanced Scorecard (wie im BIX-WB) zugeordnet und mit dem Entwurf der Norm ISO 11620 verglichen. Insgesamt fanden sich 57 Indikatoren, von denen aber nur 20 in mehreren Projekten auftreten – die Individualität ist eindeutig. Typisch ist das Beispiel der Mitarbeiterfortbildung:

- In der Norm wird die Anzahl der Teilnahmen gezählt
- BIX zählt den Zeitaufwand
- Das niederländische Projekt zählt die Kosten

Ähnlich vielseitig ist immer noch die Terminologie. Wenn z.B. SCONUL „pro Nutzer“ angibt, sind damit nicht aktive, sondern potentielle Nutzer gemeint, also die Mitglieder der Zielgruppe. Immerhin entsprechen 25 der 57 Indikatoren weitgehend der Norm 11620. Hier zeigt sich der Vorteil der Standardisierung: Die Ergebnisse werden auch über das einzelne Projekt hinaus vergleichbar.

Der Aspekt „Ressourcen“ ist mit 17 Kennzahlen vertreten. Die Rolle der Bibliothek als Arbeitsplatz wird über die Größe des Benutzungsbereichs, die verfügbaren Arbeitsplätze und Computer und natürlich die am häufigsten als Kennzahl auftauchenden Öffnungszeiten bewertet. Die Informationsversorgung (Bestände) wird über Ausgaben und Medienzahl pro Kopf sowie über Erneuerungsquote und Verfügbarkeit evaluiert. Das Angebot von Nutzerschulungen sowie die Zahl der Mitarbeiter pro Kopf tauchen nur im BIX auf.

Der Aspekt „Nutzung“ ist mit 17 Kennzahlen vertreten, sieben davon mehrfach. „Bibliotheksbesuche pro Kopf“, „Ausleihen pro Kopf“ und Marktdurchdringung (Prozentsatz der erreichten Zielgruppe) tauchen jeweils dreimal auf. Die Bestandsqualität wird außerdem über Bestandsumschlag, Nutzung der neueren Bestände (Beschaffungen der letzten fünf Jahre) oder das Verhältnis Fernleihen zu Ausleihen bewertet. Teilnahmen an Nutzerschulungen und Zahl der Auskunftsanfragen pro Kopf stehen für die Informationsdienste. Nur einmal (Schweden) wird mit den Teilnahmen an Veranstaltungen der

Bibliothek auch die kulturelle Rolle bewertet. Das schwedische und das niederländische Projekt haben auch bereits gewagt, die Nutzung elektronischer Dienste zu erfassen (Sessions im Opac oder elektronische Dienste pro Kopf, Downloads pro elektronische Zeitschrift). Für den BIX-WB ist der Indikator „Virtuelle Bibliotheksbesuche“ (website visits) vorgesehen, zu dem derzeit Tests laufen.

Dass der Aspekt „Effizienz“ mit 15 Indikatoren vertreten ist, entspricht den Forderungen nach einem Nachweis über die Mittelverwendung. Fünf der Kennzahlen tauchen mehrfach auf, am häufigsten „Kosten (Bibliotheksausgaben) pro Nutzer“. „Kosten pro Besuch“ ist nur im BIX-ÖB vertreten. Die Ausgaben für die Informationsversorgung werden zu Entleihungen, zu den Personalausgaben und zu den Gesamtausgaben der Bibliothek in Beziehung gesetzt. Für den Aspekt „Mitarbeiter“ stehen die Personalausgaben pro Kopf der Zielgruppe, die Mitarbeiterstunden pro Öffnungsstunde und die Verteilung der Mitarbeiter auf Hintergrunddienste (Medienbearbeitung, Verwaltung) und Benutzung. Die Effizienz der Prozesse wird nach Schnelligkeit (Beschaffung, Medienbearbeitung, Dokumentlieferung aus Magazinen, Fernleihe) und Korrektheit (Aufstellung im Regal, positiv erledigte Fernleihen) bewertet. BIX-WB und das niederländische Projekt evaluieren die Mitarbeiterproduktivität am Beispiel der Medienbearbeitung.

Der Aspekt „Entwicklung“ (Sind ausreichend Potentiale für notwendige Entwicklungen vorhanden?) ist erst seit der Nutzung der Balanced Scorecard in Projekten der Leistungsmessung vertreten, aber in Zeiten ständiger Veränderung von besonderer Bedeutung. Dass es schwierig ist, hierfür Kennzahlen zu finden, liegt auf der Hand. Nur im BIX und dem niederländischen Projekt sind hierfür Kennzahlen zu finden (insgesamt acht). Entwicklungspotential wird einerseits über den Stand der elektronischen Dienste definiert (Ausgaben für elektronische Bestände, Personaleinsatz für elektronische Dienste), andererseits über Personalentwicklung und -motivation (Zeitaufwand und Ausgaben für Fortbildung pro Mitarbeiter, Verfügbarkeit und Fluktuation der Mitarbeiter). Schließlich wird die Fähigkeit der Bibliothek bewertet, einen ausreichenden Mittelanteil von ihrer Institution zu erzielen und zusätzliche Mittel zu gewinnen. Abgesehen von der Fortbildungsquote, die für die „lernende Bibliothek“ steht, werden gerade die Indikatoren des Aspekts „Entwicklung“ häufig angezweifelt, treffen aber in

3 z.B. University of Virginia Library, USA <http://www.lib.virginia.edu/bsc/index.html>; Deakin University, Australia <http://www.ifla.org/IV/ifla68/papers/123-106e.pdf>; Königliche Bibliothek Kopenhagen (LIBER Quarterly 14,1, 2004) <http://liber.library.uu.nl/>; University of Pretoria, Südafrika <http://www.librijournal.org/pdf/2000-3pp202-209.pdf>

4 Australian Public Libraries Comparative Report 1998-2004 <http://www.casl.org.au/papers.and.publications.cfm>

5 <http://www.biblioteksforeningen.org/>

6 UK Higher Education Management Statistics 2003-2004. Sconul 2005

7 <http://www.ukb.nl/benchmark.htm>

internationalen Diskussionen stets auf großes Interesse. Insgesamt sind in den verglichenen Projekten neuere Bibliotheksdienste – elektronische Dienste und Medien, Schulungen zur Informationskompetenz – noch recht wenig vertreten. Das liegt bei elektronischen Diensten sicher an der Schwierigkeit, verlässliche und vergleichbare Daten, zum Beispiel zur Nutzung, zu gewinnen. Der BIX-WB bringt immerhin bereits zwei Kennzahlen zu elektronischen Diensten, eine dritte soll hinzukommen. Kennzahlen zu Einfluss und Wirkung von Bibliotheken sind in den Projekten noch nicht enthalten – wenn man nicht aus hohen Nutzungszahlen auf Wirkung schließen möchte. Solche Kennzahlen sind allerdings weitgehend erst im Teststadium, noch nicht normiert und, da qualitative Methoden angewendet werden müssen, schlecht vergleichbar.

### Löst die Nutzerbefragung alle Probleme?

Einzelne Leistungsindikatoren werden immer wieder in ihrem Aussagewert kritisiert, vor allem natürlich die in Vergleichsprojekten. Angeführt wird z.B.,

- dass die Werte durch die Bibliotheken beeinflusst werden können (Beispiel: Kurze Ausleihfristen oder effizientere Rückforderungsroutinen führen zu höheren Ausleihzahlen),
- dass von der Bibliothek nicht beeinflussbare Umstände die Werte verändern können (höhere Finanzierung, ein Bibliotheksneubau),
- dass Indikatoren „politisch“ ausgerichtet seien (z.B. wenn die Förderung elektronischer Dienste als positiv gesehen wird).

Die Lösung liegt nahe, Nutzerbefragungen als einzige Methode der Evaluierung einzusetzen. Schließlich sollen sich ja Bibliotheksleistungen am Nutzerbedarf orientieren. Weltweit meistgenutzt ist derzeit die Methode des von der Association of Research Libraries (ARL) initiierten Projekts LibQUAL.<sup>8</sup> LibQUAL bietet ein Befragungsinstrument an, das die konkrete Erfahrung mit dem erwarteten Qualitätslevel vergleicht und dadurch Lücken in den Leistungen ermittelt. Nutzerbefragun-

gen als einzige Methode der Evaluierung einzusetzen scheint aber fraglich:

- Die Einschätzung eines Services wird von früheren Erfahrungen beeinflusst. Gute Erfahrung kann zu hohen Erwartungen führen, schlechte zu geringer Erwartung und damit höherer Bewertung einer Dienstleistung.
- Benutzer können auch ohne greifbaren Nutzen zufrieden sein.
- Die Antworten können von einer gewissen Loyalität zur eigenen Bibliothek beeinflusst sein. So erzielen z.B. Aspekte wie „Freundlichkeit des Personals“ meist hohe Werte.
- Die Antworten sind oft „Momentaufnahmen“, beeinflusst von aktuellen Ereignissen, die außerhalb des Einflusses der Bibliothek liegen (z. B. Baulärm, Wetterverhältnisse).

Die Nutzermeinung gibt sozusagen die „gefühlte Qualität“ wieder. Eine eher objektive Messung kann andere Werte ergeben. Wenn etwa 30 % bestellter Titel bereits verliehen sind, werden Benutzer die es als „sehr häufig“ werten, obgleich sie bei 70 % ihrer Bestellungen sofort erfolgreich sind. Sowohl der BIX-WB wie das schwedische Benchmarking-Projekt nehmen eine Zufriedenheitsumfrage in die Liste der Indikatoren auf. Das niederländische Modell fügt der Bewertung nach Indikatoren zwei Nutzerbefragungen hinzu. Für den BIX ist eine Online-Umfrage im Test, die von allen Teilnehmern genutzt werden könnte. Im Ranking könnte dann allerdings nur eine einzige Zahl für die allgemeine Zufriedenheitsquote mit der Bibliothek erscheinen. Für Management-Entscheidungen sind eher die Einzelaussagen wichtig, die den Leistungsindikatoren gegenübergestellt werden können (z.B. Zufriedenheit mit den Öffnungszeiten).

### Lohnt sich das?

Als positive Ergebnisse gemeinsamen Benchmarkings werden gesehen:

- **Bibliotheksintern:**
  - Die Möglichkeit, durch Vergleich den Stand der eigenen Leistung besser bestimmen und Probleme aufdecken zu können.
  - Größeres Verständnis für Evaluierung und Controlling innerhalb der Bibliothek.
- **Extern:**
  - Höhere Transparenz von Mittelverwendung und Leistungsqualität.
  - Imagegewinn durch die Präsentation der Bibliotheksarbeit (Bibliotheken werden sichtbar in ihrer Leistung).

– Das Gemeinschaftsprojekt bietet eine sozusagen neutrale Plattform gegenüber der Selbstdarstellung einer Bibliothek.

### Als Probleme werden genannt:

- Wenn die gemeldeten Daten nicht überprüft werden, entstehen unstimmmige Vergleichswerte.
- Die Veröffentlichung schlechter, aber auch herausragender Werte kann für die einzelne Bibliothek nachteilig sein.
- Bibliotheken könnten über dem Bemühen, bei dem Benchmarking bessere Werte zu erzielen, andere wichtige Veränderungen zurückstellen.
- Freiwillige Teilnahme führt zu Problemen, wenn die Teilnehmer häufig wechseln und dadurch die Werte nicht über Jahre vergleichbar sind.

Eindeutig ist, dass gemeinsames Benchmarking keine fertigen Lösungen anbietet, sondern auf Probleme, Entwicklungen und mögliche Verbesserungen hinweist. Management-Entscheidungen bedürfen im Einzelfall zusätzlicher Evaluierung.

Für das niederländische Projekt fassten Laeven und Smit die Ergebnisse nach drei Jahren Laufzeit so zusammen: „We conclude by saying 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“.<sup>9</sup>

Aber geht es nicht wirklich einfacher als mit einem ganzen System von Kennzahlen?

Peter Brophy (CERLIM, Manchester) schlug in einer Diskussion zu später Stunde vor, den Erfolg einer Bibliothek mit einer einzigen Kennzahl zu messen: Prozentsatz der Benutzer, die beim Verlassen des Bibliotheksgebäudes lächeln (durch Kamera erfasst). Gegenargument: Durch das außen herrschende Wetter äußerst beeinflussbar.

„The search for an ultimate measure of benefit may be illusory.“<sup>10</sup>

„The search for an ultimate measure of benefit may be illusory.“

<sup>8</sup> [www.libqual.org](http://www.libqual.org)

<sup>9</sup> Laeven, Hubert, Anja Smit: A project to benchmark university libraries in The Netherlands. In: Library Management 24 (2003), S. 291-304. [www.ukb.nl/benchmark.htm](http://www.ukb.nl/benchmark.htm)

<sup>10</sup> Revill, Don: Performance Measures for Academic Libraries. In Kent, E. (Ed.) Encyclopedia of Library and Information Science, Dekker: New York, Basel (1990), S. 333