

# Von „Publish or Perish“ zu „Get Evaluated or Perish“

Alice Keller, Zürich

Noch nie in der Geschichte der Menschheit gab es so viele Evaluationen, d.h. Bewertungen, Einschätzungen oder Beurteilungen von Unternehmen, Einrichtungen, Produkten oder Einzelpersonen. Wo man nur hinschaut, wird evaluiert. Alle wollen wissen, wie gut sie sind! Dass die Evaluation auch schlecht ausfallen könnte, verdrängt man vorerst.

So hat das schlechte Abschneiden der 15-Jährigen beim Lesen, Rechnen und in den Naturwissenschaften im Schülervergleich Pisa ganz Deutschland schockiert. Gleichzeitig dürften solche Resultate auch dazu führen, dass eine Reihe von dringenden Reformen zu einer allgemeinen Verbesserung der Situation führen.

Das Thema Evaluation interessiert gleichermaßen sowohl in der Arbeitswelt als auch im Privatleben. Schon die Kinder sollen oder wollen evaluiert werden. So durfte ich als Kind in den Sommerferien an einem Schwimm-Wettkampf teilnehmen. Mit grossem Stolz konnte ich am Ziel einen Preis entgegen nehmen. Irritiert hat mich einzig, dass meine ältere Schwester, die etwas schneller gewesen war als ich, keinen Preis erhielt. Erst später im Leben habe ich gelernt, was ein Trostpreis ist!

Aber zurück zur Welt der Erwachsenen: Ein Blick auf die Webseite der Stiftung Warentest zeigt, dass es kaum Produkte gibt, die nicht auf ihre Qualität oder auf ihr Kosten-Nutzen-Verhältnis hin geprüft werden können. Auch die Webseite von Öko-Test hilft Ihnen beim Entscheid, wie Sie Ihren Haushalt umweltgerecht gestalten können. Von Dampfreinigern wird allgemein abgeraten, Brotbackautomaten hingegen sind erlaubt. Die Planung Ihrer Freizeit müssen Sie auch nicht dem Zufall überlassen: Beim Kinobesuch können Sie sich auf die Oscars, beim Buchkauf auf den Amazon.de-Verkaufsrang verlassen. Zur sinnvollen Gestaltung Ihres Privatlebens gibt es inzwischen so viele Hilfsmittel, dass der Spruch: „Wer die Wahl hat, hat die Qual“, eigentlich nicht mehr zutreffend sein müsste.

An wissenschaftlichen Bibliotheken bzw. an Universitäten gewinnen zunehmend Produkte an Bedeutung, die bei der Evaluation von Forschung und Wissenschaft unterstüt-

zend wirken. Schon seit langer Zeit bekannt ist die Bibliographie Science Citation Index mit dem stark umstrittenen, jedoch häufig zitierten Impact Factor. (Wenn Sie Ihr Leseverhalten ausschliesslich nach dem ISI Impact Factor richten würden, dürften Sie vermutlich meine Kolumnen nie lesen!) Während dem es hier um die Evaluation von Zeitschriften geht, stellt die neue biomedizinische Datenbank Faculty of 1000 den Wissenschaftler bzw. Forscher in den Mittelpunkt. In der Datenbank Faculty of 1000, abgekürzt F1000, werden ausgewählte Aufsätze von Experten des Fachgebiets evaluiert und mit „recommended“, „must read“ oder „exceptional“ benotet. Insgesamt sind es ca. 1.400 Wissenschaftlerinnen und Wissenschaftler, die regelmässig (d.h. monatlich 2-4) Artikel prüfen und ihre qualifizierte Einschätzung dazu abgeben.

Eines der Ziele von F1000 ist es, den einzelnen Aufsatz anstatt der ganzen Zeitschrift in den Vordergrund zu stellen. So sollen auch Arbeiten in weniger renommierten (und teuren) Zeitschriften eine Chance haben, gute Noten zu erzielen. In eigenen Worten verfolgt die Datenbank folgendes Ziel: „Highlight papers based on their scientific merit rather than the journal in which they appear, and systematically organize and evaluate scientific information.“

Das Produkt konnte an manchen Universitäten erfolgreich eingeführt werden und scheint auch ein (neue geschaffenes?) Bedürfnis von Wissenschaftlerinnen und Wissenschaftlern zu erfüllen. Dieser Erfolgskurs ist sicher u.a. auf das sehr aggressive Verkaufsverhalten des Verlags zurück zu führen. Selbstverständlich findet auch F1000 seine Kritiker!

\* Ist das Urteil der Begutachter wirklich unparteiisch?

\* In einigen Untersuchungsgebieten sind es so wenige Experten, dass bereits eine gute Note für eine Empfehlung reicht.

\* Interdisziplinäre Forschungsgebiete werden kaum berücksichtigt.

\* Ist es realistisch, dass die 1.400 Begutachter über längere Zeit diese Aufgabe gewissenhaft ausüben? Diese Personen sind erfahrungsgemäss bereits mit anderen Auf-

gaben stark überlastet. Das ganze Modell ist in hohen Masse abhängig von der regelmässigen Mitarbeit solcher Koryphäen.

\* Obwohl der Zeitschriftentitel für die Auswahl der Aufsätze nicht ausschlaggebend sein soll, scheinen die Flaggschiffjournale regelmässig die Spitzenplätze einzunehmen. In einem Editorial in Nature Medicine wird diese Kritik wie folgt zusammengefasst: „And despite the figure quoted by F1000's operators that 70% of recommendations are for papers other than those in Nature, Science and Cell, 5 minutes searching the site suggests that many papers come from these very journals“. (Tatsächlich sind gegenwärtig von den zehn All Time Top Aufsätzen sieben aus den Zeitschriften Nature, Science oder Cell!)

Dieser Trend hin zur Evaluation einzelner Artikel heisst natürlich für Bibliotheken, dass zusätzliche Produkte gekauft bzw. lizenziert werden müssen. Es wird in Zukunft nicht mehr reichen, wenn wissenschaftliche Bibliotheken Primärliteratur und einfache bibliographische Findmittel zur Verfügung stellen. Bereits heute wird von vielen Wissenschaftlerinnen und Wissenschaftlern eine weitere Stufe der Inhaltsanalyse oder –bewertung (ranking) gewünscht. Die Datenbank Faculty of 1000 nennt sich bezeichnenderweise „next generation literature awareness tool“.

Es erstaunt natürlich nicht, dass das erste Produkt dieser neuen Generation aus dem Fachgebiet der Biomedizin stammt. Hier dürfte der globale Wettbewerb zwischen den Universitäten bzw. Forschungsgruppen am grössten sein. Das Motto der Wissenschaft lautet zukünftig nicht mehr „Publish or Perish“, sondern vielmehr „Get Evaluated or Perish“.

Obwohl es im Bibliothekswesen keinen Nobelpreis gibt, möchten auch Bibliotheken bewertet, beurteilt und benotet werden. Evaluation im Bibliothekswesen bedeutet in der Regel den Vergleich verschiedener Einrichtungen oder Dienstleistungen anhand geeigneter Kenngrössen bzw. Leistungsindikatoren.

Einen sehr interessanten Beitrag zum Thema Leistungsmessung in Bibliotheken lieferte

## Kolumne

Prof. Konrad Umlauf an der diesjährigen ASpB-Tagung in Stuttgart<sup>1</sup>. Im Gegensatz zur Situation im Profitbereich, wo sich der Erfolg vor allem aus dem Gewinn bzw. aus der Rentabilität ableiten lässt, müssen Bibliotheken andere Indikatoren zur Leistungsmessung suchen.

K. Umlauf hat in seiner Studie neun ausgewählte Handbücher zur Leistungsmessung analysiert und hierbei insgesamt – sage und schreibe – 200 verschiedene Leistungsindikatoren für Bibliotheken gefunden! Diese Vielzahl von Indikatoren ist einerseits darauf zurück zu führen, dass die Terminologie uneinheitlich verwendet wird. Andererseits besitzen Bibliotheken als Horte des Sammeln unzählige Objekte, die man zählen, messen oder beurteilen kann.

Die Evaluation von K. Umlauf zeigt, dass von den knapp 200 Indikatoren nur elf mindestens in vier Instrumentarien zur Leistungsmessung zugleich erscheinen. Nur gar drei Indikatoren finden in sechs der neun Instrumentarien zugleich Anwendung. Kein Indikator kommt häufiger vor. Die Frage von K. Umlauf scheint berechtigt zu sein: Ist dieses Resultat ein Hinweis auf unausgereifte Konzepte oder auf Beliebigkeit? Eine interessant Frage, zur der ich keine Antwort finde.

Dass ein allgemeiner Vergleich verschiede-

ner wissenschaftlicher Bibliotheken schwierig ist, zeigt auch die Arbeit von G. Reichmann<sup>2</sup>. Der Autor geht hier der Frage nach, ob signifikante Leistungsunterschiede zwischen Universitätsbibliotheken des deutsch- und englischsprachigen Raumes existieren. Gleichzeitig untersucht er die qualitativen Unterschiede zwischen kleinen und grossen Universitätsbibliotheken.

Anhand dieser Studie erkennt man eindeutig, wie schwierig bzw. gefährlich es ist, Kennzahlen aus deutschsprachigen und angelsächsischen Universitätsbibliotheken zu vergleichen. So vergleicht er beispielsweise die Zahl der Serials an US-amerikanischen Bibliotheken mit der Zahl der laufenden Zeitschriftentitel an deutschsprachigen Hochschulen. Mit dieser Zählweise nimmt die University of Illinois at Urbana-Champaign mit über 90.000(!) Zeitschriftentiteln den Spitzenplatz ein. Erst nach über dreissig Universitätsbibliotheken aus dem englischsprachigen Raum folgt mit der UB Graz (knapp 17.000 laufende Titel) die erste Einrichtung aus dem deutschsprachigen Raum. Prof. W. Neubauer schreibt in seiner sehr kritischen Rezension in der Zeitschrift *Information: Wissenschaft & Praxis* dazu: „Wie bereits oben erwähnt, so zeigt sich auch hier in aller Deutlichkeit, dass der Versuch eines qualifizierten

Vergleichs von Universitätsbibliotheken des deutschsprachigen und angloamerikanischen Raumes einfach scheitern muss, wenn man mit dem Untersuchungsgegenstand nicht angemessen vertraut ist. Jedem bibliothekarischen Anfänger ist bekannt, dass „serials“ eben nicht Zeitschriften im engeren (deutschsprachigen) Sinne sind, sondern dass hier alles mitgezählt wird, was sich als fortlaufende Publikation präsentiert.“

Prof. Dr. Alice Keller  
ETH-Bibliothek  
Rämistrasse 101  
CH-8092 Zürich, Schweiz  
Tel. ++41 1 632 67 20  
Fax. ++41 1 632 13 57  
E-Mail: [alice.keller@library.ethz.ch](mailto:alice.keller@library.ethz.ch)

<sup>1</sup> Umlauf, Konrad: Kritische Analyse der Leistungsmessung von Bibliotheken und Informationseinrichtungen. - Vortrag gehalten auf der ASpB-Tagung in Stuttgart am 9. April 2003.

<sup>2</sup> Reichmann, Gerhard. Universitätsbibliotheken im Vergleich - eine internationale Querschnittsuntersuchung Gerhard Reichmann. Wiesbaden: Deutscher Universitäts-Verlag; 2001.