

## Wissenschaftsevaluation mittels Datenbanken – methodisch einwandfrei?

Als Maß für die Produktivität und den Einfluß von Forschern, wissenschaftlichen Einrichtungen und Fachbereichen dienen häufig anhand von Publikations- und Zitationsanalysen erstellte Ranglisten. Doch nach welchen Kriterien sind die in elektronischen Fachdatenbanken gespeicherten Informationen auszuwerten, um ein einigermaßen zutreffendes Abbild der Forschungsleistung zu erhalten?

Von Wolfgang G. Stock

In weit mehr als tausend wissenschaftlichen und technischen Datenbanken unterschiedlicher Spezialisierung und Größe ist das Wissen unserer Zeit interaktiv abfragbar gespeichert. Die Retrieval-Software gestattet außer zielgenauem Suchen nach einzelnen Literaturstellen auch sogenannte informetrische Recherchen – statistische Charakterisierungen von gewissen Mengen von Datensätzen. Untersuchungsgegenstände der Informetrie sind die Schriften eines Autors, die Forschungen eines Instituts, eines Landes, einer Sprachregion, eines Themengebietes oder einer Disziplin. Die informetrischen Maße werden so zu Indikatoren wissenschaftlicher Produktivität und deren Wirkung.

Doch trotz der weiten Einsatzgebiete dieser Indikatoren in der Wissenschaftsevaluation sind deren methodologische Probleme alles andere als gelöst. Kritiker wie Wolfgang Glänzel von der Ungarischen Akademie der Wissenschaften in Budapest und Urs Schöpfli vom Berliner Max-Planck-Institut für Wissenschaftsgeschichte werfen der Indikatorenforschung vor, sie diene vorwiegend den kurzfristigen Interessen der Wissenschaftspolitik und -planung und reduziere ihren wissenschaftlichen Gehalt auf die bloße Präsentation von Datensätzen – Grund genug also, sich mit den Problemen der Wissenschaftsevaluation und ihrer Methoden, der Publikations-, Themen- und Zitationsanalysen, zu befassen.

### *Publikationen als Indikatoren*

Forschungsleistung schlägt sich in der Regel in Texten nieder, in Fachartikeln oder -büchern, in offengelegten Patentschriften oder in Gutachten. Datenbanken sammeln nach jeweils eigenen Auswertungsregeln denjenigen Teil der wissenschaftlichen und technischen Texte, der als dokumentationswürdig eingestuft wird.

Je nach Erfassungsregelwerk werden Autoren- und Institutsnamen unterschiedlich angesetzt, wobei auch Tippfehler nicht ausbleiben. Von wenigen Ausnahmen abgesehen werden die Einträge nach Themen selektiert und nicht nach Autoren oder Institutionen. So ergibt sich bereits daraus ein Problem für die Evaluation von Forschungsleistungen: Vollständigkeit ist auch durch Ansprechen diverser Datenbanken nicht zu garantieren; es bleibt also in der Regel ungewiß, ob die Menge der Literatur von oder zu einem Wissenschaftler oder einem Institut komplett ist. Lediglich Patentliteratur ist nahezu vollständig in Datenbanken erfaßt.

Publikationsraten erfassen die Anzahl von Veröffentlichungen. Doch welche Einheit soll man wählen? Gewisse Studien zählen bei der Summenbildung jede Publikation als Einheit – gleich, ob es sich um eine kurze Notiz oder ein umfangreiches Lehrbuch handelt. Um dieses Problem zu umgehen, verwenden

andere Autoren nach Dokumententyp gewichtete Publikationsraten. Eine Analyse der Anglistik beispielsweise vergibt einen Gewichtungswert von 50 für eine Monographie, von 10 für einen Aufsatz und von 1 für eine Rezension. Andere Gewichtungen sind durchaus vorstellbar und auch bereits benutzt worden.

Um Werke mit mehreren Autoren zu behandeln, bieten sich zwei Methoden an. Man kann für jeden beteiligten Autor die Publikation mit „1“ zählen, doch dürfte dies seinen Anteil an der Forschungsleistung nicht adäquat widerspiegeln. Zudem ist die Summenbildung der Aggregate Institut, Hochschule oder Land problematisch, weil das Aufaddieren der einzelnen individuellen Publikationsraten einen Wert größer als 100 Prozent ergibt. Nach einem anderen Verfahren werden deshalb die Anteile gezählt. Beispielsweise bekäme jeder Coautor eines Werkes mit insgesamt drei Verfassern für diese Publikation den Wert 1/3 zugeordnet. Dies erscheint fairer, sagt jedoch auch nichts über den faktischen Anteil der Mitarbeit am Forschungsprojekt aus. Weil multiple Autorschenschaft in der heutigen Forschung die Regel ist, differieren die nach der Anteils- beziehungsweise nach der Gesamtzählung bestimmten Publikationsraten sehr stark.

Auch Unterschiede in den Publikationsorganen sind zu berücksichtigen. Dazu errechnet man den sogenannten *journal impact*. Die ursprüngliche Methode geht auf Eugene Garfield, einen Pionier der Zitationsindexierung und Leiter des Institute for Scientific Information (ISI) in Philadelphia (US-Bundesstaat Pennsylvania), zurück; sie setzt Publikationen einer Zeitschrift der beiden zurückliegenden Jahre  $t-1$  und  $t-2$  in Relation zu den Zitationen, die diese Veröffentlichungen im Jahr  $t$  erhalten haben. Doch begünstigt eine solche Berechnungsweise, die nur ein oder zwei Jahre alte Zitationen berücksichtigt, hochaktuelle Zeitschriften, deren Informationen gewissermaßen direkt konsumiert werden, und benachteiligt all jene, deren Zitation sich über einen längeren Zeitraum erstreckt.

Ein weiterer Aspekt der Gewichtung von Periodika ist die sogenannte *Freifahrten-Hypothese*. Ihr zufolge werden Artikel, die in einer prominenten Zeitschrift erscheinen, häufiger zitiert als solche in anderen Erscheinungsorganen. Doch läßt sich ein hoher *journal impact* nicht unbedingt auf den Einzelfall übertragen, wie empirische Untersuchungen belegen. Vielmehr ist es so, daß einige

bahnbrechende Arbeiten in den prominenten Zeitschriften eine große Resonanz in der Fachwelt hervorrufen und häufig zitiert werden. Andere Beiträge in denselben Organen hingegen werden kaum oder gar nicht beachtet. Wie fast überall bei informatrischen Verteilungen liegt auch hier eine typische linkschiefe Verteilung vor, die das Berechnen eines aussagefähigen Mittelwertes unmöglich macht. Auf diese Problematik der Wissenschaftsevaluation hat der Mannheimer Wissenschaftsforscher Hans-Dieter Daniel bereits vor Jahren hingewiesen.

Bei der Bewertung von Publikationsanalysen ist zu beachten, daß sie eben nur Veröffentlichungen zählen und nicht alle Forschungsleistungen. Diese können sich durchaus völlig anders manifestieren: in Artefakten beispielsweise wie Bauwerken oder Computerprogrammen. Insbesondere Technikwissenschaftler weisen ihre Forschungs- und Entwicklungsleistungen auf diese Weise nach und nicht oder höchstens rudimentär durch Literatur.

Ferner können Publikationsanalysen nicht die wissenschaftlichen Inhalte der

Veröffentlichungen erfassen. Hierzu sind themenanalytische informatrische Methoden erforderlich. Dabei handelt es sich um ein Bündel statistischer Verfahren, die auf inhaltsabbildende Angaben in Datensätzen elektronischer Datenbanken angewandt werden. Mit ihnen läßt sich etwa zeigen, ob ein Forscher stets zum gleichen Thema publiziert oder ob seine Forschungsgebiete variieren (Bild 1). Weil themenanalytische Studien jedoch aufwendig sind, werden sie noch viel zu selten durchgeführt.

#### *Zitationen und Thematisierungen als Wirkungsindikatoren*

Die bisher beschriebenen Methoden sind zwar Indikatoren für Forschungsleistungen, sagen allerdings nichts über Erfolg oder Mißerfolg der Publikationen aus. Gerade dies ist aber sowohl für den Fortschritt der Wissenschaft als auch für die Anwendung wichtig. Der Wiener Wissenschaftstheoretiker Erhard Oeser schreibt hierzu: „Ein wissenschaftlich relevanter Erkenntnisprozeß muß das ihm zugehörige Informationssystem be-

einflussen. Das heißt: Er muß im System eine Entscheidung hervorrufen. Kann er das nicht, dann ist er entweder überflüssig oder nicht zugehörig.“ Wirkungen lassen sich durch Thematisierungen (im Text) und Zitationen (in den Fußnoten oder Literaturverzeichnissen) ablesen.

Für Zitationsanalysen gibt es drei unterschiedliche Ansätze:

– Zitationsraten sind Auszählungen der Häufigkeit von Zitationen, angefangen bei Zitationen genau eines Werkes und aggregiert – je nach Forschungsinteresse – beispielsweise für einen bestimmten Autor, eine wissenschaftliche Institution oder ein Land. Unterschieden werden dabei absolute Zitationsraten von Relativwerten, unter anderem Zitationen pro Jahr, pro zitierfähiger Publikation, pro publizierter Seite oder pro Zeichen.

– Co-Zitations-Analysen suchen wissenschaftliche Werke in einen Zusammenhang zu bringen. Zwei Zitationen gelten als verknüpft, als co-zitiert, wenn sie in den zitierenden Arbeiten gemeinsam auftauchen. Ein solcher Ansatz bietet sich insbesondere dann an, wenn sehr viele zitierende Arbeiten gemeinsame Zitate enthalten. Zur Darstellung faßt man die Co-Zitations-Paare in Cluster zusammen, die in der Wissenschaftsevaluation als Indikator für gemeinsame Forschungsaktivitäten gelten. Dabei bezieht sich der Clusterkern auf die (co-) zitierten Publikationen, die sogenannte Forschungsfront auf die zitierenden und somit neueren Publikationen. Zentraler Aspekt der Evaluation einzelner Forschungsarbeiten beziehungsweise der Forschung von Instituten, Universitäten oder Ländern ist deren Vorkommen entweder als Zitationen in den Clusterkernen (dann handelt es sich um Arbeiten, die eine Forschungsrichtung mitbestimmen) oder als den Clusterkern zitierende Werke in der Forschungsfront (dann liegen Arbeiten vor, die Teil neuartiger Forschungen sind). Ist ein Autor im ersten Fall Auslöser einer wissenschaftlichen Neuerung, so ist er im zweiten zumindest in Forschungsfronten präsent.

– Zitationen weisen auf einen Informationsfluß vom Zitierten zum Zitierenden hin. Durch Analyse dieser Informationsvermittlung ergibt sich ein Indikator für Wissenschafts- beziehungsweise Techniktransfer. Der Informationsfluß vermag somit die Stellung eines Wissenschaftlers oder einer forschenden Institution im internationalen Wissenschafts- und Techniktransfer abzubilden.

Indes – beim Anwenden von Zitationsanalysen ist Vorsicht geboten, denn der Verweis auf eine andere Publikation

Die häufigsten Forschungsthemen Vittorio Benussis:

1: Gestalt, 2: geometrisch-optische Täuschung, 3: Suggestion, 4: Farbe, 5: Zöllnersche Figur, 6: Täuschungsgröße, 7: Vorstellung, 8: Zeit, 9: Psychologie, 10: Scheinbewegung

Themennetz um „Gestalt“ im Lebenswerk Benussis:



**Bild 1:** Das Lebenswerk des österreichisch-italienischen Psychologen Vittorio Benussi (1878 bis 1927) zeigt in der informatrischen Themenanalyse Schwergewichte auf der Gestaltforschung, der geometrisch-optischen Täuschung sowie der Suggestion (oben). Für Benussis Hauptforschungsgebiet wurden die thematischen Bezüge in Form eines semantischen Netzes analysiert.

Wie aus dem Teilcluster rechts im Netz hervorgeht, experimentierte Benussi in der Gestaltforschung bei der Analyse geometrisch-optischer Täuschungen mit der Winkelgröße der Müller-Lyerschen Figur. Durch solche Themenanalysen, hervorgegangen aus einer auf Vollständigkeit angelegten Datenbank, sind aussagekräftige szientometrische Ergebnisse ableitbar.

als Quelle kann unterschiedliche Bedeutung haben.

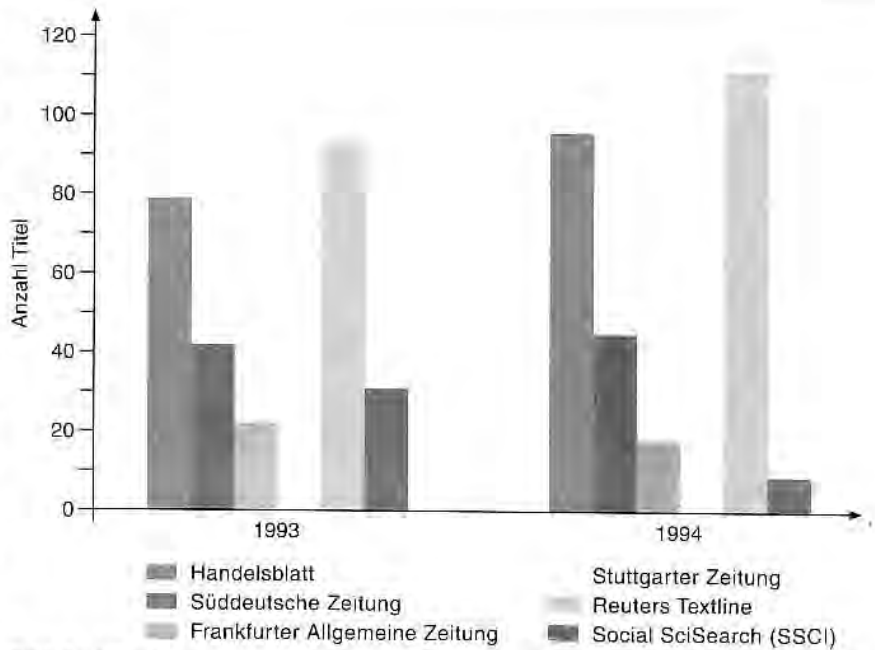
Nicht immer ist der zitierte Artikel als Ganzes gemeint. Oft wird nur ein bestimmter Abschnitt als relevant eingestuft. Selbst eine einzige Phrase wie etwa eine besonders gut gelungene Formulierung oder der Name für eine Methode kann Anlaß für das Zitat gewesen sein. Andererseits mag die zitierte Quelle nur ein Teil eines größeren Komplexes sein, etwa wenn Autoren Motive zitieren (zum Beispiel: Sexualität bei Freud) oder das Lebenswerk eines Autors. Der Informationswissenschaftler Blaise Cronin von der Universität von Indiana in Bloomington hat dafür den Begriff „Zitationslage“ eingeführt.

Selbstzitationen zeigen an, auf welche früheren Schriften ein Autor aufbaut. Beläßt man sie bei der Zählung der Zitationsraten in der Grundgesamtheit, so vermag ein Autor seine Stellung in einer Rangliste zu verbessern, indem er auch thematisch unspezifische eigene Werke ausgiebig zitiert. Selbstzitationen prinzipiell auszuschließen würde wiederum bedeuten, gewisse Informationskanäle auszublenden, weil es doch gerade aufschlußreich sein kann zu beobachten, wie Ideen sich im Laufe eines Forscherlebens weiterentwickeln. Zudem ist stets auch der Fall möglich, daß der Autor außer eigenen Werken nichts Zitierfähiges vorfindet – etwa weil er wissenschaftliches Neuland betreten hat oder, im anderen Extrem, sich noch als einziger in einem Gebiet abmüht, von dem sich andere längst abgewandt haben.

Auch sogenannte Zitationskartelle – Gruppen von Wissenschaftlern, die sich gegenseitig zitieren – sind schwierig zu bewerten. Im positiven Falle könnte dieses Verhalten auf gemeinsame Forschungsgebiete hinweisen; doch wäre auch denkbar, daß die Forscher damit geschickt ihre Zitationsrate zu steigern suchen, weil solche Kartelle im Gegensatz zu gewöhnlichen Selbstzitationen bei Analysen nur schwer zu identifizieren sind.

Mengxiong Liu von der San José State University (Kalifornien) analysierte in einer Übersichtsarbeit die Praxis der Zitation und insbesondere die Motive der zitierenden Autoren. Sie fand, daß die Zitierenden in der Regel überzeugen wollen und dabei Literaturstellen nennen, mit denen sich die eigene Meinung unterstützen läßt.

Nach alledem schiene es recht fragwürdig, wenn jede Nennung eines Werkes in einem Artikel jeweils gleichgewichtet mit „1“ gezählt würde. Hinzu kommt das



**Bild 2:** Mitarbeiter des ifo Instituts für Wirtschaftsforschung in München wurden 1993 und 1994 nur 40mal in der Datenbank Social SciSearch des Institute for Scientific Information (ISI) genannt. Demnach würde dieses Institut keine allzu große Wirkung auf die internationale Wissenschaftlergemeinschaft haben. Doch diese

Schlußfolgerung wäre verfehlt: Mit jeweils rund 100 Titelnennungen in „Reuters Textline“, fast ebensovielen in der führenden deutschen Wirtschaftszeitung „Handelsblatt“ und mehreren Dutzend Schlagzeilen in weiteren überregionalen Tageszeitungen ist die Wirkung auf Wirtschaft, Politik und Öffentlichkeit nicht zu unterschätzen.

bereits erwähnte Problem der Mehrfachautoren. Soll das Zitat nur dem erstgenannten Autor, allen Autoren jeweils mit „1“ oder allen Autoren anteilmäßig zugeordnet werden? Soll man Zitationen aus Zeitschriften mit unterschiedlichem *impact* auch unterschiedlich gewichten? Und welche Rolle spielen Zitationen in Büchern?

Sind ferner die Indikatorwerte für unterschiedliche Wissenschaftsdisziplinen miteinander vergleichbar? Das Zitationsverhalten bezüglich der Publikationskanäle ist äußerst verschieden. Nur fünf Prozent aller Zitationen von Chemikern betreffen Monographien, während 75 Prozent der Soziologen gerade diese bevorzugen. Verallgemeinernd läßt sich sagen, daß Naturwissenschaftler eher Artikel zitieren und Geistes- beziehungsweise Sozialwissenschaftler eher Bücher.

Die Publikations- und Zitationsgewohnheiten in den verschiedenen Wissenschaftsdisziplinen unterscheiden sich des weiteren durch ihre jeweils durchschnittlichen Publikationsraten, die Anzahl der Zitationen pro Artikel, die Länge ihrer Aufsätze und das Alter der zitierten Arbeiten. Infolgedessen ist bei fächerübergreifenden Vergleichen der absoluten Zahlen von Indikatoren stets größte Vorsicht geboten.

Die meisten Zitationsuntersuchungen stützen sich auf Material des ISI, das – abgesehen von einigen Patenzitatenbanken – auf dem Gebiet der wissenschaftlichen Zitationen eine Monopolstellung innehat. Dessen Nachweise nennen grundsätzlich nur den jeweils ersten Autor des zitierten Werkes. In der Analyse entfallen deshalb die Namen eventueller weiterer Autoren, oder sie müssen mühevoll nachrecherchiert werden. Weil zudem anstelle des oder der Vornamen nur Initialen angegeben sind, ist es schwierig, Personen gleichen Namens zu unterscheiden.

Von den derzeit weltweit etwa 100 000 vertriebenen wissenschaftlichen Zeitschriften wertet das ISI nur etwa 5000 aus; Bücher werden mit Ausnahme von Kongreßberichten nicht erfaßt. Die Auswahl der Periodika richtet sich nach einer Rangfolge von Zeitschriften, die sich aufgrund einer nach Garfield benannten Regel ergibt. Dieses Gesetz der Konzentration besagt, daß eine recht geringe Anzahl von Zeitschriften im gesamten Wissenschaftsbereich große Zahlen von Zitationen auf sich vereinigt, während auf die überwiegende Menge der Periodika wenig oder gar keine entfallen. Für jedes Wissenschaftsgebiet sind Kontingente definiert, die mit den

---

jeweils meistzitierten Zeitschriften dieser Disziplin gefüllt werden. Folglich haben in ihrer quantitativen Verbreitung weniger bedeutende theoretische Ansätze und wenig gesprochene Sprachen kaum eine Chance, vom ISI bearbeitet zu werden.

Zitationsanalysen reichen nur so weit, wie es üblich ist zu zitieren. Das bedeutet, daß sich die Wirkung einer Forschungsleistung außerhalb des engen Bereichs der Wissenschaft grundsätzlich nicht mittels Zitationsanalyse erfassen läßt. Dazu ist ein weiterer Indikator erforderlich. In einem Forschungsprojekt am ifo Institut für Wirtschaftsforschung arbeiteten wir erfolgreich mit der Thematisierung in Zeitungen und Agenturmeldungen. Solche Untersuchungen stützen sich auf elektronische Volltextdatenbanken, in denen jedes Wort abfragbar ist. Auf diese Weise läßt sich ein lückenloses Bild aller Pressemeldungen zu einem Autor oder Institut erhalten. Zudem sind Pressedatenbanken aktuell, denn ein Text ist in der Regel spätestens am Tag nach seinem Erscheinen darin gespeichert.

Wie relevant die Thematisierung als Wirkungsindikator ist, belegen empirische Studien. Mit ihr lassen sich beispielsweise Fehlinterpretationen vermeiden, die zwangsläufig aufträten, wenn man sich ausschließlich auf die in Zitationsdatenbanken gespeicherten Informationen verließ. So zeigt sich erst durch Einbeziehen der Thematisierungen des Namens eines betrachteten Instituts in den Presseschlagzeilen das Ausmaß seiner Wirkung.

Die Aussichten für die Wissenschaftsevaluation erscheinen letztlich als optimistisch. Die bisherige eindimensionale Betrachtung von Publikations- und Zitationsraten ist zu erweitern. Werden die ergänzenden Möglichkeiten informatrischer Verfahren voll ausgeschöpft, indem man insbesondere Themenanalysen durchführt, und die methodischen Probleme ernst genommen, so können elektronische Datenbanken wertvolles Rohmaterial für die Wissenschaftsevaluation bereitstellen.

---

Dr. Stock ist seit April 1995 Professor für Informationswissenschaft und Wissenschaftstheorie an der Fachhochschule Köln. Zuvor leitete er das Informationszentrum des ifo Instituts für Wirtschaftsforschung in München, wo auch das hier vorgestellte Projekt entstanden ist.

---