

Rafael Ball

Bibliometrie im Zeitalter von Open und Big Data

Das Ende des
klassischen Indikatorenkanons



Impressum

© Dinges & Frick Verlag, Wiesbaden 2015

www.b-i-t-online.de

ISBN 978-3-934997-72-1

Rafael Ball

Bibliometrie im Zeitalter von Open und Big Data

Das Ende des klassischen Indikatorenkanons

Inhalt

1	Vorwort	7
2	Einleitung und Geschichte	15
3	Bibliometrische Verfahren und ihre Grundlagen	27
3.1	Die Menge von wissenschaftlichen Publikationen: Outputanalyse	31
3.2	Die Messung der Wahrnehmung: Resonanzanalyse	32
3.3	Die Bestentabellen: Rankings und Benchmarking auf der Basis von Bibliometrie	35
3.4	Bibliometrie als Trendscout: Trendanalysen und Vorhersagen	41
3.5	Neue und alternative bibliometrische Verfahren: Altmetrics	43
3.6	Von der Kausalität zur Korrelation: Big Data in der Bibliometrie	50
4	Bibliometrie in den Geistes- und Sozialwissenschaften: Besondere Formen und Verfahren	57
4.1	Einleitung	57
4.2	Publikations- und Zitationskulturen in den Geistes- und Sozialwissenschaften	59
	Zitationskulturen	66
4.3	Formen des Wissenschaftsoutputs ausgewählter Disziplinen	68
5	Bibliometrische Indikatoren	73
5.1	Die Anzahl wissenschaftlicher Veröffentlichungen	74
5.2	Zitierhäufigkeit und Zitierrete	78
5.3	Der Hirsch-Index	84
5.4	Der Impaktfaktor	90
5.5	Alternative Metriken	99

6 Die Datengrundlage	106
6.1 Web of Science und Science Citation Index von Thomson Reuters	108
6.2 SCOPUS von Elsevier	114
6.3 Google Scholar	116
7 Zusammenfassung und Ausblick	120
Zitierte und weiterführende Literatur	125
Internetadressen	143
Anhang	145
Glossar	146
Abbildungen	153

1 Vorwort

Die Bibliometrie wird im Jahr 2016, also im nächsten Jahr, 100 Jahre alt. Sie ist keine neue Erfindung des Medienunternehmens Thomson Reuters, das mit den einschlägigen Datenbanken, etwa dem Science Citation Index einen riesigen Umsatz generiert und viel Geld damit verdient. Sie ist auch keine Erfindung des Medienkonzerns Elsevier, der mit der Datenbank SCOPUS der einzige ernst zu nehmende Wettbewerber zum Science Citation Index darstellt und heute mit einer ganzen Reihe von Wissenschaftsevaluations-Produkten die Hochschulen flutet.

Bibliometrie ist vielmehr entstanden aus der Idee, Bibliothekare bei ihrer eigentlichen Arbeit, der Literatursauswahl, zu unterstützen und das Bestandsmanagement zu optimieren. Das war die Grundidee der ersten bibliometrischen Analysen, und das war auch der Ansatz des amerikanischen Chemikers Eugene Garfield, der in den 50er-Jahren des 20. Jahrhunderts begonnen hatte, Zeitschriftenartikel systematisch auf die benutzte und zitierte Literatur hin auszuwerten. Und auch seine Ergebnisse sollten der Beratung und Unterstützung von Bibliothekaren dienen.

In den ersten 50 Jahren des 20. Jahrhunderts waren also Fragen nach der Art und Häufigkeit von Wissenschaftspublikation keine wirklich wesentlichen Themen, ebenso wenig Fragen danach, wie Wissenschaft funktioniert und Publikationen entstehen. Sie waren daher noch viel weniger Gegenstand quantitativer Untersuchungsmethoden für den wissenschaftlichen Output, wie es heute die Bibliometrie darstellt.

Ohnehin war der Output der Wissenschaftler in dieser Zeit noch völlig unwichtig. Wer ein guter Professor war, hatte viele Studenten, er bekam dann eine erkleckliche Summe an so genannten Hörergeldern und lebte, zumindest in Deutschland,

mit einem Einkommen, das zu dem höchsten ein Prozent der Gesellschaft zählte.¹ Der Direktor einer großen Bank verdiente zu Beginn des 20. Jahrhunderts rund ein Drittel weniger als ein Professor an einer Universität.

Der schriftliche Forschungsoutput in Form von Büchern und Zeitschriftenartikeln oder in Form von Vorträgen war zu diesem Zeitpunkt und vor dem geschilderten Hintergrund noch nicht von Belang. Nur wer etwas zu sagen hatte, fand Gehör: Er schrieb Bücher, machte Vorträge und publizierte in den sich erst langsam etablierenden Zeitschriften wissenschaftliche Beiträge.

Wer das aber nicht tat, hatte auch kein Problem, und er hatte vor allem auch keine Konsequenzen zu befürchten. Zielvereinbarungen über eine Mindestanzahl von wissenschaftlichen Publikationen pro Jahr, wie es heute an vielen medizinisch-naturwissenschaftlichen Instituten längst üblich ist, waren damals jenseits des Vorstellbaren.

So ist die Bibliometrie als Methode zur Bewertung der wissenschaftlichen Leistungen von Personen und Institutionen denn auch ein Kind der Massenwissenschaft geworden, wie sie nach dem Zweiten Weltkrieg entstanden ist. Die Anzahl der Wissenschaftler nahm dramatisch zu, Disziplinen wurden aufgesplittet, und damit stieg natürlich der Bedarf an notwendigen spezialisierten Publikationsorganen wie Zeitschriften, Buchreihen und Konferenzen. Auch gab es immer mehr Verlage und immer mehr Veröffentlichungen. So war es nur eine Frage der Zeit, bis die Notwendigkeit entstanden war, den schriftlichen Output der Wissenschaftler und ihrer Institutionen zu erfassen und zu bewerten.

Bibliometrie misst zunächst die Menge der publizierten Schriften einer Person. Dies ist die leichteste Übung, da sie noch keiner Annahmen bedarf und noch unmittelbare Aus-

1 FAZ, Seite N4, Mittwoch, 09.04.2014, NR. 84, Forschung und Lehre

sagen produziert, etwa der Art, dass ein Wissenschaftler im Verlauf seiner Karriere 200 Zeitschriftenbeiträge und 5 Bücher veröffentlicht hat. Schwieriger wird die Angelegenheit, wenn man versucht, die Qualität und Bedeutung der Aufsätze und Bücher zu bestimmen. Hier ging die Bibliometrie einen einfachen und bis heute praktisch angewandten Weg: Eine Veröffentlichung ist umso wichtiger, je mehr sie wahrgenommen wird. Für die Quantifizierung der Wahrnehmung haben die Bibliometriker die Zitation als Indikator gewählt: Eine Publikation, die in anderen Veröffentlichungen oft zitiert wird, ist eine wichtige Publikation – eine Veröffentlichung, die nicht oder nur selten zitiert wird, ist eine weniger wichtige.

Eine Veröffentlichung wird publiziert, andere Wissenschaftler zitieren diese Veröffentlichung, der ganze Vorgang wird gemessen und aufgeschrieben. Die nahezu unendliche Variation der Indikatoren, die heute in der kaum mehr zu überschauenden Literatur der Bibliometrie diskutiert wird, ändert nichts an der Tatsache, dass das Prinzip von damals bis heute gleich geblieben ist. Die Fundamentalannahme der Bibliometrie bedeutet grundsätzlich, dass ein Beitrag umso wichtiger ist, je öfter er zitiert wird. Wer dieser Grundannahme nicht folgt, kann Bibliometrie nicht nutzen. Wer sie begründet ablehnt, wird aus der Bibliometrie keine Erkenntnisse generieren können.

Zudem sind umfassende Datenbanken (die früher als gedruckte Verzeichnisse herausgegeben wurden) nötig, die die jeweiligen Zitierungen nachweisen, damit sie ausgewertet werden können. Der schon erwähnte Amerikaner Eugene Garfield hat in den 1950er-Jahren mit der systematischen Auswertung von Zeitschriftenartikeln begonnen. Bis heute ist daraus eine gewaltige Datensammlung geworden, auf der ein Großteil bibliometrischer Analysen basiert.

Die klassische Bibliometrie fußt seit 50 Jahren auf der Datenbank des Science Citation Index, seiner Erweiterungen und

Features. Trotzdem sind das zugrunde liegende Messprinzip und vor allem die Annahmen und Rückschlüsse aus den erhobenen Daten bis heute gleich geblieben. Auch der nahezu einzige kommerzielle Konkurrent, die Datenbank SCOPUS von Elsevier, macht im Prinzip nichts anderes.

Auch inhaltlich und logisch basieren diese Systeme auf der Denkweise des cleveren Gründers des Web of Science bzw. des Science Citation Index Eugene Garfield.

Die Zeiten der Bibliometrie, die auf der Fundamentalannahme der einfachen Zuordnung von Veröffentlichung und Zitation beruht, sind noch lange nicht vorbei. Doch es kommen neue Methoden hinzu, die die alten eines Tages sicher ablösen werden: „Usage Metrics“ etwa haben einen grundsätzlich anderen Ansatz als die klassische Zitationsbibliometrie und basieren nicht mehr auf dem Fundamentalansatz, der fast 100 Jahre gültig war. Usage Metrics lösen zwar das Problem der Bedeutungsbeurteilung eines wissenschaftlichen Beitrags nicht grundsätzlich – niemand kann durch mittelbare Methoden bestimmen, wie wichtig ein Beitrag ist, wie oft er gelesen wird und wie viel Erkenntnis der Leser tatsächlich aus diesem Beitrag gewinnt. Aber es gelingt durch Usage Metrics eine deutliche Annäherung an diese Grundfrage, die letztlich die eigentliche Kernfrage der Bibliometrie ist.

Usage Metrics eröffnen erstmals die Chance, die Bedeutung von wissenschaftlichen Veröffentlichungen nicht mehr nur über die ausschließlich indirekte Vermittlung der Zitierate zu bestimmen, sondern erlauben ein Verhältnis von Nutzung und Bedeutung herzustellen.

Diese Korrelation kann allerdings noch weiter optimiert werden. Dank Usage Metrics hat man z.B. die Möglichkeit, den Download einer Publikation nachzuweisen. Es ergeben sich weitere Anwendungen, auch die Zeit des Bearbeitens und die Verweildauer des Nutzers im Dokument, die Art seiner Bearbeitung, etwa durch Markierungen oder Kopieren, fest-

zuhalten und statistisch auszuwerten. Auch die Weiterleitung eines Dokuments oder das Teilen mit anderen in sozialen Medien kann eine Aussage über die Bedeutung eines wissenschaftlichen Beitrags liefern.

Libmetrics (Library Metrics) wiederum stellen einen Zusammenhang her zwischen der Bedeutung eines wissenschaftlichen Beitrags oder Buches und der Verfügbarkeit und Nutzung in der Bibliothek. Dies geschieht etwa durch die Messung der Häufigkeit der Beschaffung oder der Ausleihe dieser Literatur in der Bibliothek. Die Durchdringungstiefe von Bibliotheksbeständen, mit einem Buch etwa, kann eine Aussage enthalten, wie relevant ein solches Werk eingeschätzt wird. Die Auflage des Buches oder der generierte Umsatz durch den Buchhandel lassen Rückschlüsse auf die Attraktivität eines Werkes zu. Libmetrics stecken allerdings noch tief in den Kinderschuhen. Die Veröffentlichung von wissenschaftlichen Beiträgen in elektronischer Form, sei es als Zeitschriftenbeitrag, elektronisches Buch, Blog, Chat, Website-Beitrag oder als multimedialer Beitrag in unbestimmtem Medienformat, ist inzwischen die etablierte Form wissenschaftlicher Kommunikation und der Verbreitung von Erkenntnissen.

Das freie Internet, die Entwicklung der Wissenschaftskommunikation und die sozialen Netzwerke eröffnen neben neuen Möglichkeiten auch andere Notwendigkeiten, die Messung und Bestimmung des wissenschaftlichen Outputs neu zu denken.

Neben den bereits erwähnten Optionen der „Altmetrics“ werden die Nutzung und Kombination aller freien Netzdaten und die Anwendung von Big-Data-Technologien auf das System der Veröffentlichungen und ihrer Messung zu neuen Einsichten führen: „Weil das Internet solche Informationen überall auf der Welt verfügbar macht und die Technologie zu unserer zweiten Natur geworden ist, kann die digitale Öffent-

lichkeit in Echtzeit von Dingen, Menschen, Erfahrungen und Ereignissen berichten.“²

Die klassische Bibliometrie und ihre Indikatoren haben zwar noch nicht ausgedient, aber ihre Aussagen werden zunehmend auf das reduziert, was sie bedeuten: Sie bieten eine nur sehr mittelbare Annahme, dass eine oft zitierte Publikation wichtig ist.

Bereits heute ist die Hälfte aller wissenschaftlichen Beiträge der Europäischen Union (EU) frei im Netz verfügbar.³ Mit der weiteren Entwicklung von frei verfügbaren wissenschaftlichen Netzinhalten diesseits der Bezahlschranke haben sich längst neue Möglichkeiten ergeben, mehr, direkter und treffsicherer über die Bedeutung von Wissenschaftsoutput zu befinden.

Während die Menge des wissenschaftlichen Outputs explodiert, findet Wissenschaft und Forschung heute längst nicht mehr im unbeobachteten Schutzraum des wissenschaftlichen Elfenbeinturms statt, sondern auf der Bühne des gesellschaftlichen „Schlachtfeldes“ um Gelder, Ehre und Anerkennung. Ob die Wissenschaft und ihre Forscherinnen und Forscher das wahrhaben wollen oder nicht – sie konkurrieren beim Kampf um die Ressourcen der Staaten mit Schwimmbädern, Autobahnen, dem neuen europäischen Kampfhubschrauber und mit den ausufernden Sozialausgaben.

Die Gesellschaft will es sich nicht mehr länger leisten, Wissenschaft bedingungslos vor sich hin arbeiten zu lassen, sie verlangt zurecht Rechenschaft auch von jenen, die mit Steuer-

-
- 2 Bunz, M.: Die stille Revolution: wie Algorithmen Wissen, Arbeit, Öffentlichkeit und Politik verändern, ohne dabei viel Lärm zu machen. Berlin: Suhrkamp 2012, S. 145
 - 3 Archambault, E. et al.; Science-Metrix (Hrsg.): „Proportion of Open Access Peer-Reviewed Papers at the European and World Levels – 2004–2011“; August 2013, online abrufbar unter http://www.science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf

geldern Forschung betreiben, deren Sinn sich nicht erschließt. Und sie erwartet ebenfalls zurecht, wie ich meine, dass sich die Wissenschaftler messen lassen mit vergleichbaren Maßstäben, die international anerkannt und vergleichbar sind – und nicht durch die undurchdringliche Selbstbestätigung des „Inner Circles“, einer eingeschworenen Community, die sich permanent selbst bescheinigt, wie gut und hervorragend ihre Ergebnisse und wie fundamental ihre Forschungen für die Zukunft der Menschheit sind. Der Steuerzahler möchte wissen, welche Maßstäbe anzulegen sind, damit eine faire und ehrliche Messung jener Leistungen von Wissenschaft und Forschung, aber auch von Wissenschaftlerinnen und Wissenschaftlern möglich ist, die die Gesellschaft finanziert, die aber gleichzeitig auch die Exzellenz der beteiligten Personen und Einrichtungen ausmachen oder eben auch nicht.

Die Quantifizierung des wissenschaftlichen Outputs ist das Kernthema von Bibliometrie, das Messen der Leistungsperformance von Einrichtungen und Personen – mit all den Unwägbarkeiten, die in Wägemaßen so stecken.

Viele Disziplinen, etwa die Medizin, große Teile der Naturwissenschaften und auch Teile der Wirtschaftswissenschaften, verfügen längst über ein etabliertes System der Leistungsbeurteilung des wissenschaftlichen Outputs. Die Indikatoren sind weitgehend akzeptiert, was aber nicht bedeutet, dass auch in diesen Fachgebieten noch vieles verbessert werden kann, sondern dass neue Metriken durch technologischen Fortschritt (Webometrie), aber auch durch veränderte Publikationsgewohnheiten möglich werden.

Aber noch nicht jede Disziplin findet sich bereit, sich in die Karten schauen zu lassen. Die Geisteswissenschaften, Teile der Sozial- und Wirtschaftswissenschaften und auch die Rechtswissenschaft beharren noch viel zu sehr auf qualitativen Kriterien für die Bewertung von Wissenschaftsleistungen – teils aus Angst, in der subjektiven Schutzatmosphäre der

wissenschaftlichen Freiheit gestört zu werden, teils aus Unkenntnis über die Methoden der Bibliometrie, teils aus Furcht vor der Anwendung quantitativer Messgrößen, zu denen man außerhalb der Naturwissenschaft ein gespaltenes Verhältnis hat.

Allerdings muss man fair bleiben: Die Publikationskultur in den Geistes- und Sozialwissenschaften unterscheidet sich von der in der Medizin und den Naturwissenschaften deutlich, und tatsächlich ist es nicht immer einfach, ihren wissenschaftlichen Output quantitativ zu fassen. Zudem hat die Bibliometrie dieses Thema zu lange sträflich vernachlässigt. Es war einfacher, die Zitationszahlen eines biologischen Papers über die Jahre zu addieren als die Forschungsleistung eines Editionswissenschaftlers in der Germanistik angemessen zu würdigen.

Bibliometrische Fragestellungen, die von der Medizin über die Naturwissenschaften bis zu den Geistes- und Sozialwissenschaften reichen, sind komplex und umfangreich. Dazu bedarf es auch der Kenntnis und Würdigung der jeweiligen Publikationskulturen. Nur so lassen sich Hürden und Vorbehalte abbauen und Methoden entwickeln, die jeweils angemessen sind.

Rafael Ball
Zürich, im März 2015

2 Einleitung und Geschichte

„Jedermann wird benotet. Der Liebende von der Geliebten, die sich in Schweigen hüllt, der Händler durch lautstarke Kundenbeschwerden, die Medien durch die Quote, der Arzt durch Patientenzustrom, der Gewählte durch Reaktionen der Wähler.“⁴ (Michel Serres)

Von wissenschaftlichen Publikationen im heutigen Sinne spricht man nach De Solla Price erst seit den 20er-Jahren des 19. Jahrhunderts, obwohl die erste wissenschaftliche Zeitschrift *Philosophical Transactions* bereits 1665 erschienen ist.⁵ Allerdings kann von der so genannten Großforschung (big science), die sich durch das exponentielle Wachstum wissenschaftlicher Publikationen auszeichnet, erst seit der zweiten Hälfte des 20. Jahrhunderts gesprochen werden. De Solla Price war der Meinung, dass das exponentielle Wachstum der Wissenschaft im Jahr 2000 durch rund 1.000.000 wissenschaftlicher Zeitschriften und 1.000 relevanter sekundärer Informationsquellen sichtbar sein werde. Diese Zahl stimmt nicht ganz mit der tatsächlichen Anzahl wissenschaftlicher Zeitschriften überein, die im ULRICHs (Ulrichs International Periodicals Directory) auf rund 180.000 beziffert wird. Auf der Makroebene, die sich auf alle Disziplinen und Fachrichtungen bezieht und alle Länder umfasst, wächst die Anzahl der Zeitschriften um ca. 3,3% pro Jahr.⁶

-
- 4 Serres, M.: Erfindet euch neu! Eine Liebeserklärung an die vernetzte Generation. Berlin: Suhrkamp 2013, S. 49
 - 5 Price, D.J. de Solla: Little Science, Big Science. Columbia: University Press 1963
 - 6 Mabe, M.M.A.: Growth dynamics of scholarly and scientific journals. *Scientometrics* (2001) 51(1), 147–162

De Solla Price war es auch, der zwei Jahre später (1965) in einem Artikel in *Science* auf das Thema der Netzwerke zwischen Autoren hingewiesen hatte und dazu das Zitierverhalten von Wissenschaftlern untersuchte.⁷ Diese so genannten Clusteranalysen sind heute ein zeitgemäßes bibliometrisches Forschungsfeld.

Die Bewertung wissenschaftlicher Produktivität und Produktion bzw. deren an Zitaten gemessenen Resonanz ist deshalb umso notwendiger, je größer die Zahl des publizierten wissenschaftlichen Outputs ist.

In Zeiten knapper Kassen ist die leistungsorientierte Mittelvergabe dringend auf objektive Parameter der Leistung und Leistungsfähigkeit der Wissenschaft und ihrer Wissenschaftler angewiesen. Dabei wird auch die Frage, ob die Qualität wissenschaftlicher Arbeit einer letztlich quantitativen Begründung durch die mittelbaren Annahmen der Bibliometrie zugänglich ist, ebenso hinterfragt wie die bereits Realität gewordene quantitative Beurteilung von einzelnen Wissenschaftlern, Instituten, Einrichtungen oder Ländern selbst. Die Tatsache allein, dass die Anwendung bibliometrischer Methoden in der Wissenschaftsvaluation schon weit verbreitet ist, macht es notwendig, das überaus komplexe Thema für alle beteiligten Gruppen einerseits leicht verständlich und schnell lesbar, andererseits aber auch in der gebotenen Tiefe des komplexen Themas aufzuarbeiten.⁸

Forschungs- und Wissenschafts-Rankings sind populär. Es gibt keinen Universitätspräsidenten, der nicht die Position seiner Einrichtung im nationalen und internationalen Vergleich kennt (und sie erklären oder instrumentalisieren kann), und kein Politiker kommt ohne sie aus: „Wissen-

7 Price, D.J. de Solla: Networks of scientific papers. In: *Science* (1965) 149, S. 510–515

8 Ball, R.: *Bibliometrie: Einfach – verständlich – nachvollziehbar* (Praxiswissen), Berlin: De Gruyter, Saur Verlag 2013

schaft und Politik werden sich gegenseitig beeinflussen und beobachten.“⁹ Bibliometrie wird heute zunehmend verknüpft mit Leistungsmessung, Leistungsbewertung, Evaluation und wissenschaftspolitischer Steuerung.

Bibliometrie als Quantifizierung von wissenschaftlichem Output ist die Folge eines Rechtfertigungsdrucks in Wissenschaft und Forschung, der zunächst einer zweckfreien Wissenschaft im aristotelischen Sinne als Erkenntnis um der Erkenntnis Willen widerspricht. Noch heute zeugt das Recht auf Freiheit in der Forschung, das Professoren in vielen Ländern an Universitäten zugestanden wird, von diesem ursprünglichen Wissenschaftsverständnis.

In den 1960er-Jahren allerdings begann eine Änderung der öffentlichen Wahrnehmung dieser Wissenschaftsprivilegien im Forschungsumfeld öffentlich finanzierter Universitäten und Forschungszentren. Dies geschah in Folge der Demokratisierung der Gesellschaft und ihrer Universitäten.

Die Hörengelder, oder auch so genannte Kollegelder, wurden abgeschafft, und es flossen keine finanziellen Mittel mehr für die bloße Anzahl der Studierenden. Die Partizipation der Studenten begann in den 1968er-Jahren als Revolution, die den „Muff von tausend Jahren aus den Talaren“ vertreiben wollte und damit auch das Unantastbare im Status eines Professors, seiner Forschungsfreiheit und seines Status quo in der Organisation der Universität sowie der bürgerlichen Gesellschaft relativierte. Damals gab es rund 5000 Professoren in Deutschland.

Mit der Anzahl der Wissenschaftler nahm auch die Anzahl der Promotionen und Habilitationen massenhaft zu. Das schwemmte eine Menge von Wissenschaftlern an die Univer-

9 Taubert, N.: Bibliometrie in der Forschungsevaluation. Zur Konstitution und Funktionslogik wechselseitiger Beobachtung zwischen Wissenschaft und Politik. In: pub.uni-bielefeld.de/publikationsserver

sitäten und beendete mit der Massenuniversität der 1970er-Jahre den herausgehobenen Status eines Professors endgültig. Heute tragen mehr als 40.000 Personen in Deutschland den Professorentitel. Der Professor wurde praktisch zu einem normalen Mitarbeiter an der Universität. Als Rest jenes elitären Status, den ein Professor zu Beginn des 20. Jahrhunderts noch hatte, bleibt nur noch das bis heute geradezu reflexhaft eingeforderte Recht auf Freiheit von Forschung und Lehre, was mithin bedeutet, dass jeder ordentlich berufene Professor forschen und lehren kann, was ihm beliebt, und Forschungsgegenstände wählen kann, die er für sinnvoll erachtet.

„Wer die Fallhöhe ermisst, die der deutsche Professor vom Mandarin zum normalen Berufstätigen dieser Tage zurücklegte, kann die Melancholie verstehen, mit der seine Zunft noch immer den heroischen Jahren der deutschen Universität vor 1914, dem Beginn des Ersten Weltkriegs, nachhängt“, schrieb „DIE ZEIT“ in dem Beitrag „Mythos Professor“ 2011.¹⁰

Die Einführung der Leistungsmessung durch die Quantifizierung des Wissenschaftsoutputs, etwa durch Bibliometrie, ist heute nahezu die einzige Möglichkeit, auf die Performance der Wissenschaftler einzuwirken, ist doch das Recht auf freie Wahl der Forschungsgegenstände noch immer ein Teil des (deutschen) Berufungssystems.

Oft wird daher die Quantifizierung von Leistungen im Wissenschaftsumfeld (mit bibliometrischen Methoden) als ein der Wissenschaft unangemessener Zwang empfunden, genauso wie die jüngste Regelung des Landes Baden-Württemberg, das seinen Wissenschaftlern vorschreibt, die mit Steuergeldern erworbenen Forschungsergebnisse im Open Access Modus zu publizieren und damit der Gesellschaft wieder kostenlos zur Verfügung zu stellen.¹¹

10 <http://www.zeit.de/2011/30/C-Mythos-Professor/seite-2>

11 http://de.wikipedia.org/wiki/Heidelberger_Appell

Die Einführung von Bibliometrie an Universitäten und bei Wissenschaftlerinnen und Wissenschaftlern und ihre Akzeptanz sind deshalb nicht nur ein Methoden- und Verständnisproblem oder ein Problem der wagen Datengrundlage und der unterschiedlichen Publikationskulturen, sondern gleichzeitig ein psychologisches, weil es vielen Wissenschaftlern ihrem Status nicht angemessen erscheint, ihre Performance quantitativ nachweisen zu müssen.

Gleichzeitig basiert das 21. Jahrhundert auf einer Allgegenwärtigkeit von Daten. Das Motto der CEBIT 2014 prägte deshalb auch das Kunstwort „Datability“. Die Verfügbarkeit von Daten im Netz und die Präsenz persönlicher Daten in sozialen Medien lassen aber auch Wissenschaftler in dieser neuen Dimension der „Datability“ denken und handeln. Negativ gewendet, beschreibt das Malirsch so:

„Der Mensch erkennt sich als Produkt, das verkauft werden muss, eine wandelnde Reklame, ein Verwalter des eigenen Lebenslaufs und der Entrepreneur der eigenen Möglichkeiten (...).“¹²

So gefasst liegt es im Interesse eines Wissenschaftlers, seine eigenen Leistungen auch quantitativ nachzuweisen, ja sogar anzupreisen. Dies kann dann sogar dazu führen, dass man aus Eigeninteresse seine Erkenntnisse zunächst zurückhält, um sie zu einem geeigneten Zeitpunkt nur für sich geltend zu machen. Die heutigen Möglichkeiten der Kooperation, des Teilens und des gemeinsamen Arbeitens am Erkenntnisfortschritt stehen aber noch immer in einem seltsamen Missverhältnis zu den individuellen Belohnungs- und Karrieresystemen in der öffentlichen Wissenschaft und Forschung.

„Viele Wissenschaftler haben die Ironie aufgezeigt, dass gerade jetzt, zu dem Zeitpunkt in der Geschichte, da wir die Technologien

12 Malirsch, M.: Big Data: Büchse der Pandora. Remscheid: Re Di Roma Verlag 2013, S. 79

haben, die eine weltweite Verfügbarkeit von wissenschaftlichen Daten und verteilte Verarbeitung dieser ermöglichen, wo Zusammenarbeit vertieft werden und Entdeckungen beschleunigt werden können, dass genau zu diesem Zeitpunkt wir unsere Zeit darauf verwenden, eben diese Daten verschlossen zu halten und dadurch die Anwendung ebenso fortschrittlicher Technologien zu ihrer Erschließung verhindern.“¹³

Für die Auswertung des wissenschaftlichen Outputs und ihren Einfluss auf die Entwicklung eines Fachgebietes werden grundsätzlich zwei Methoden eingesetzt: die qualitative und die quantitative. Zum einen ist dies die wissenschaftliche Begutachtung (Peer Review), zum anderen die bibliometrische Analyse.

Die Peer-Review-Methode kann im Unterschied zu bibliometrischen Methoden kaum durch objektive Merkmale bestimmt werden. Es handelt sich dabei um Einschätzungen kompetenter Fachexperten, die oft als subjektiv bezeichnet werden. Zumindest lässt sie sich nicht quantifizieren.

Bibliometrische Analysen hingegen gewähren Einsicht in alle wichtigen Komponenten der Wissenschaft auf der Ebene von Makroforschungen:

- die Struktur wissenschaftlicher Aktivitäten in einzelnen Disziplinen auf nationaler Ebene,
- die wissenschaftliche Produktivität, aufgeschlüsselt nach Ländern,
- der Einfluss einzelner Länder oder Regionen auf bestimmte Wissensgebiete,
- internationale und regionale Zusammenarbeit,
- die Informiertheit von Wissenschaftlern über Entwicklungen einzelner Wissenschaftsbereiche in der unmittelbaren oder weiteren Umgebung,

13 Wilbanks, J., Executive Director, Science Commons http://de.wikipedia.org/wiki/Open_data

- die Benutzung formaler Kommunikationskanäle in einem Land,
- das wissenschaftliche Produkt und sein Einfluss,
- institutionelle Zusammenarbeit.

Objekte dieser Forschungen sind Produzenten von Publikationen (individuelle Autoren, Teams, Einrichtungen, Länder und Regionen), die Publikationen selbst (Zeitschriften, Artikel, sekundäre Informationsquellen) sowie deren deskriptive Eigenschaften und Zitatensanalysen, die wissenschaftliche Kommunikationsprozesse erschließen.¹⁴

Bibliometrie ist ein Instrument zur Ermittlung objektiver Publikationsdaten, die oft als Leistungsdaten Verwendung finden. Sie kann dabei helfen, die erwähnten Aufgaben zu erfüllen und die Daten dabei hochmathematisch und statistisch betrachten oder für den Nicht-Mathematiker in ihren Grundzügen versteh- und durchschaubar machen. Widerstand und Bedenken gegen die Ermittlung von Daten zur Forschungsleistung kommen vor allem von Menschen, die die Methode nicht verstehen, Angst vor ihr haben und ihr Manipulation unterstellen. Deshalb ist es wichtig, Bibliometrie versteh- und nachvollziehbar zu machen, aber auch gleichzeitig ihre Grenzen aufzuzeigen und neue Alternativen zu benennen.

Der moderne Begriff von Bibliometrie ist noch sehr jung. Er wurde 1969 von Pritchard als „statistical bibliography“ geprägt¹⁵, wohingegen bereits 1934 der Begriff in einem anderen Zusammenhang gebraucht wurde.

14 Moed, H.F., De Bruin, R.E., Nederhof, A.J., Van Raan, A.F.J., Tijssen, R.J.W.: State of the art of bibliometric macro-indicators. An overview of demand and supply. In: *Science* (1992) 149, S. 510–515

15 Pritchard, A.: Statistical bibliography or bibliometrics? *Journal of Documentation* (1969) 25, S. 348–349: „Therefore it is suggested that a better name for this subject [statistical bibliography] is bibliometrics, i.e. the application of mathematics and statistical methods to books and other media of communication.“

Gorraiz definiert sie heute als „Anwendung mathematischer und statistischer Methoden zur Erklärung der Prozesse der schriftlichen Mitteilungen“¹⁶.

Die Inhalte von Bibliometrie sind dagegen viel älter, allerdings in einer Anwendung, wie sie heute kaum noch existiert: als bibliographische Statistik zur Untersuchung von Veröffentlichungen zu speziellen thematischen Fragestellungen. Die Ergebnisse dieser Untersuchungen dienten dabei ausschließlich der inhaltlichen Bewertung zu konkreten Fragestellungen nach Sachthemen. Damit zeigt sich, dass die eigentliche und ursprüngliche Anwendung von Bibliometrie eher bibliothekarisch war.

So ist uns die erste bibliometrische Analyse von Cole and Eales überliefert. Die Autoren haben im Jahre 1917 untersucht, welche Bücher zur menschlichen Anatomie im Zeitraum von 1550 bis 1860 erschienen waren.¹⁷ Der Beitrag gilt als erste bibliometrische Analyse, obwohl er noch keine Zitate verwendete. Ziel der Arbeit war es, inhaltliche Schwerpunkte und Schwankungen im Umfang der Publikationen zu diesem Thema im Zeitraum festzustellen. Sie war damit eine reine Output-, aber keine Wahrnehmungsanalyse.

Die erste bibliometrische Arbeit, die Zitationen untersucht hat, ist die Arbeit von Gross und Gross aus dem Jahre 1927.¹⁸ Sie analysierten die in Fußnoten gemachten Zitate im Fachgebiet Chemie und konnten so ein Ranking der wichtigen, weil oft zitierten chemischen Zeitschriften ihrer Zeit aufstellen. Diese Information nutzte zum einen der Fachcommunity zur Beurteilung der wichtigen Publikationsorgane und

16 Gorraiz, J. Szientometrie: Zitatelanalyse. 2004, zitiert nach: zbp.univie.ac.at/gj/citation/skriptum2neu.htm.

17 Cole, F.J., Eales, N.B.: The history of comparative anatomy. Part I: A statistical analysis of the literature. *Science Progress* (1917) 11, 578–596

18 Gross, P.L.K., Gross, E.M.: College libraries and chemical education. *Science* (1927) 66, 385–389

entspricht damit der Grundidee eines Zeitschriftenrankings und des heute noch so wichtigen Impaktfaktors. Zum anderen waren Gross und Gross Bibliothekare und verfolgten mit ihrer Untersuchung die Absicht, eine Entscheidungshilfe bei der Beschaffung von Zeitschriften durch Bibliotheken zu liefern. Sie stellten bei ihrer Analyse eine ungleichmäßige Verteilung der Zitationen auf die verschiedenen Zeitschriften fest und lieferten damit die Basis des von Bradford im Jahre 1934 entwickelten Gesetzes, wonach die entscheidenden wissenschaftlichen Publikationen auf wenige Kernzeitschriften konzentriert sind.¹⁹ Brian C. Vickery hat diese „schiefe Verteilung“ 1948 als „Bradford’s Law of Scattering“ mathematisch beschrieben.²⁰

Schon 1926 forschte der amerikanische Mathematiker Alfred James Lotka über die Produktivität der Wissenschaftler und beschrieb den Zusammenhang von Autor und Publikation, wonach der Publikationsausstoß umgekehrt proportional zur Anzahl der Wissenschaftler in einem Fach war. Das heißt: Nur wenige Autoren haben viele Publikationen und viele Autoren wenige. Dieser Zusammenhang wird auch als „Lotkas Gesetz“ bezeichnet.²¹

In den USA begann sich der Chemiker Eugen Garfield für das Thema zu interessieren und eine Sammlung von Publikationsdaten und ihren Zitaten anzulegen. Bis dahin hatte er die *Current Contents* herausgegeben, ein Verzeichnis von Inhalten wichtiger Zeitschriften. In einem Zeitschriftenartikel schlug

-
- 19 Bradford, S.: Sources of Information on Specific Subjects. In: Engineering (1934) 137, S. 85f
20 Vickery, B.C.: Bradford’s Law of Scattering. Journal of Documentation (1948) 4, 198–203
21 Lotka, Alfred J.: The frequency distribution of scientific productivity. Journal of the Washington Academy of Sciences (1926) 16: 317–323

er bereits 1955 die systematische Erfassung und Auswertung von Zitationen in wissenschaftlichen Publikationen vor.²²

Da er für seinen Index keine Unterstützung der öffentlichen Hand erhielt, gründete er das Institut of Scientific Information (ISI) und verkaufte seine Datensammlung seit 1963 auf kommerziellem Wege.

Garfield hat mit der Etablierung des weltweit erfolgreichen Science Citation Index eine bleibende Institution geschaffen. Sein ursprüngliches Anliegen allerdings war zunächst die Unterstützung der Bibliothekare bei der Kaufentscheidung von Literatur, insbesondere von Zeitschriften. So war der gedruckte Science Citation Index auch angelegt. Eine Verwendung seiner Indikatoren zur Bewertung von Personen und Institutionen, insbesondere des berühmtesten, des Impaktfaktors (Journal Impact factor), lehnt Garfield bis heute ab.

Bis heute gilt der Science Citation Index, der längst als hochkomplexe elektronische Online-Datenbank verfügbar ist, als internationaler Benchmark für die Wissenschaftsevaluation. Erst in den 90er-Jahren des 20. Jahrhunderts entwickelte sich ein Wettbewerbsprodukt, die Datenbank SCOPUS des Wissenschaftsverlages und Medienunternehmens Elsevier. Auch die Entwicklung des Internet trug zu einem weiteren Wettbewerb im Bereich bibliometrischer Daten für Wissenschaft und Forschung bei, etwa mit der Etablierung der Datenbank Google Scholar durch den Internetkonzern Google.

Eine weitere Entwicklung hat die Bibliometrie durch die Omnipräsenz des Internet und der freien Netzinhalte auch in der Wissenschaft genommen. Neben die Zählung von Veröffentlichungen und das Analysieren von Zitaten treten nun quantitative Parameter, die das Indikatorenspektrum (z.B. die

22 Garfield, E.: Citation Indexes for Science. A New Dimension in Documentation through Association of Ideas. In: Science (1955) 122(3159), S. 108–111

Anzahl der Downloads, die Verweildauer beim Dokument, die Priorisierung der In- und Outlinks), zugleich aber auch die auswertbaren Quellen erweitern.

Grundsätzlich korrespondiert Bibliometrie immer mit der Entwicklung des Wissens und seiner Kommunikation. Die klassische Vorstellung, dass Wissen in distinkten Schritten voranschreitet, die manifest werden durch konkrete abgeschlossene Publikationen, hat sich spätestens mit Beginn des 21. Jahrhunderts aufgelöst. Mit der kontinuierlichen, fluiden Wissenschaftskommunikation von liquiden Dokumenten und unabgeschlossenen Wissensportalen ist (wieder) klar geworden, dass sich auch das Wissen per se nicht in stabilen kleinen Wahrheitspaketen, die als Veröffentlichung daherkommen, schrittweise weiterentwickelt.

„Im Westen haben wir uns Wissen über ein paar Jahrtausende hinweg als ein System stabiler und konsistenter Wahrheiten vorgestellt. Kann es sein, dass uns das mehr über die Grenzen der Medien des Wissens verrät als über das Wissen selbst? Wenn Wissen kommuniziert und konserviert wird, indem man es mit Tinte auf Papier schreibt, dann ist Wissen eben das, was es durch institutionelle Filter schafft und sich nicht verändert. Das neue Medium des Wissens ist aber weniger ein System zur Veröffentlichung von Aufsätzen oder Büchern, sondern eine vernetzte Öffentlichkeit. Vielleicht können wir mithilfe der Data Commens neues Wissen produzieren, allerdings wird dieses Wissen dann eher die Form einer permanenten Diskussion annehmen, innerhalb deren es einmal hierhin gezerrt wird und einmal dorthin. So sieht Wissen im Zeitalter des Netzes aus. Es ist nie wirklich stabil, es ist nie vollständig aufgeschrieben und es ist nie endgültig fertig.“²³

23 Weinberger, D.: Die digitale Glaskugel. In: Big Data. Das neue Versprechen der Allwissenheit. Redaktion Heinrich Geiselberger, Tobias Moorstedt. Berlin: Suhrkamp 2013, S. 237

Gegenstand bibliometrischer Messungen ist damit nun nicht mehr nur traditionell bei Verlagen erschienene Literatur, sondern im Internet (frei) zugängliches wissenschaftliches Material, wie Open-Access- Publikationen, Inhalte von Fach- und institutionellen Repositorien oder persönliche und institutionelle Websites.

Die Webometrie erweitert nicht nur den strukturellen Begriff der Bibliometrie, sondern auch deren Einsatzspektrum. Gleichzeitig zeigt sich daran, wie komplex und vielfältig die Messung des wissenschaftlichen Outputs von Personen und Institutionen mittlerweile geworden ist. Die Analyse von großen Datenmengen („Big Data“) als neue Herausforderung, oder besser: als neue Chance ermöglicht eine immer detailliertere Erfassung und Analyse von wissenschaftlichen Ergebnissen.

Rafael Ball ist promovierter Biologe, Wissenschaftshistoriker und Philosoph und arbeitet aktuell als Direktor der ETH-Bibliothek in Zürich. Er ist Autor und Herausgeber zahlreicher Bücher und Publikationen, engagierter Redner und Lehrbeauftragter an Hochschulen. Seine Interessenschwerpunkte sind Wissensorganisation, Medienwandel und Bibliotheken im Internetzeitalter.

Die Bibliometrie wird demnächst 100 Jahre alt. Sie ist entstanden aus der Idee, Bibliothekare bei ihrer eigentlichen Arbeit, der Literatursauswahl, zu unterstützen und das Bestandsmanagement zu optimieren. Heute meint Bibliometrie die Quantifizierung des wissenschaftlichen Outputs, das Messen der Leistungsperformance von Einrichtungen und Personen mit all den Unwägbarkeiten, die in Wägemaßen stecken.

Durch das freie Internet und die sozialen Netzwerke haben die klassische Bibliometrie und ihre Indikatoren zwar noch nicht ausgedient, aber ihre Aussagen werden zunehmend auf das reduziert, was sie bedeuten: Eine nur sehr mittelbare Annahme, dass eine oft zitierte Publikation eine wichtige ist. Alternative Metriken erlauben die Bewertung von neuen Formen der Wissenschaftskommunikation jenseits von Zeitschriftenbeiträgen und Büchern. Sie bedeuten zugleich eine Demokratisierung in der Vermessung der Wissenschaft durch die Emanzipation von den kommerziellen monopolartigen Zitierdatenbanken am Markt.

Verlag Dinges & Frick GmbH, Wiesbaden

ISBN 978-3-934997-72-1, € 24,50

BAND 56 der Reihe b.i.t.online-Innovativ