

Zur metrischen Evaluierung von Publikationsleistungen

Sabine Zauchner, 2010

Department für Interaktive Medien und Bildungstechnologien
Donau-Universität Krems

In den letzten Jahren hat sich im Hinblick auf Zitierindizes – metrische Maße zur Evaluierung der Publikationsleistung von Wissenschaftlerinnen und Wissenschaftlern - eine rege und kontrovers geführte Diskussion entwickelt. Neuere Indizes wie h- oder g-Index ergänzen die in manchen Disziplinen sehr gebräuchlichen Impact Faktoren bzw. legen in der Hoffnung auf eine größere Aussagekraft den Fokus auf andere Berechnungsweisen und Datengrundlagen. In der Folge soll – ohne Anspruch auf Vollständigkeit - ein kurzer Überblick über derartige Indizes gegeben werden sowie auf deren augenscheinliche Problemfelder eingegangen werden.

1. Zitierindizes

1.1 Impact Factor und Science Impact Index

Der Impact Faktor (exakt: „Journal Impact Factor“) wird vor allem in den Naturwissenschaften und der Medizin verwendet und misst, wie häufig ein Artikel aus einer spezifischen Zeitschrift von anderen Zeitschriften in Relation zur Gesamtzahl der dort veröffentlichten Artikel zitiert wird. Je höher der Impact Faktor, desto angesehenere ist eine Fachzeitschrift und die wissenschaftliche Reputation der dort publizierenden Wissenschaftler/innen. Als Datenquelle fungiert das ISI Web of Knowledge, das "Web of Science", ein kostenpflichtiges Angebot das auf die Online-Zitationsdatenbanken "Science Citation Index" und "Social Sciences Citation Index" rekurriert. Heute wird die Datenbank von Thomson Reuters betrieben (<http://sub3.isiknowledge.com/>), die im Journal Citation Report jährlich die Journal Impact Faktoren für ca. 8.000 Fachzeitschriften veröffentlichen.

Der Impact Faktor hat in Disziplinen weniger Bedeutung, die ein anderes Publikationsverhalten aufweisen, als die Naturwissenschaften und die Medizin. Während der wissenschaftliche Austausch in diesen Fachgebieten vorwiegend über Zeitschriftenartikel stattfindet, variiert die Bedeutung von Zeitschriften-Artikeln und Konferenzbeiträgen oder aber auch Monografien in Abhängigkeit vom jeweiligen Fachgebiet. In den o.g. Zitationsdatenbanken werden diese Unterschiede jedoch nicht berücksichtigt. So fehlen Artikel aus Conference Proceedings beinahe vollständig, auch Monografien sind im Web of Science unterrepräsentiert.

Eine Annäherung an dieses Problem stellt ein zusätzlicher Indikator für individuelle Forschungsleistungen dar, der „Science Impact Index“. Auch bei diesem Indikator handelt es sich um eine Zitationsrate. Diese entspricht der Anzahl der Wissenschaftler/innen, die die Arbeiten eines Autors/einer Autorin innerhalb eines Jahres in einer Zeitschrift mit Impact-Faktor zitieren. Dabei werden auch Bücher, Buchbeiträge und Veröffentlichungen in Zeitschriften ohne Impact-Faktor erfasst, wenn sie in Impact Faktor-Zeitschriften zitiert werden.

Der Impact Faktor eignet sich grundsätzlich nicht, um große Fachdisziplinen, die sich durch viele Forscherinnen bzw. Forscher und Publikationsorgane und somit höhere Zitierfrequenzen auszeichnen, mit kleineren Disziplinen zu vergleichen. Daher gilt für

einen möglichst objektiven Einsatz des Impact Faktors, dass nur Leistungen innerhalb einer Disziplin verglichen werden sollten.

1.2 h- und g-Index

Auch der h-Index (Hirsch-Index) basiert auf Zitationen der Publikationen einer Autorin bzw. eines Autors hat jedoch gegenüber dem Impact Faktor oder einer reinen Auflistung der Zitationshäufigkeit eines Autors bzw. einer Autorin den Vorteil, dass die Zitationen einer einzelnen, häufig zitierten Veröffentlichung keinen Einfluss auf den Index haben. Der g-Index wiederum nimmt die Vorteile des h-Index auf und ergänzt diese um die positive Berücksichtigung herausragender Arbeiten.

In der Folge wird die Berechnung dieser Indices beschrieben. Ausgangspunkt der Berechnungen ist eine Auflistung der Publikationen eines Autors/einer Autorin nach der Anzahl der Zitierungen.

Beispiel:

Anzahl von Publikationen einer Autorin und Zitierungen der Publikationen:

Artikel 1: 75 Zitierungen
Artikel 2: 25 Zitierungen
Artikel 3: 15 Zitierungen
Artikel 4: 15 Zitierungen
Artikel 5: 6 Zitierungen
Artikel 6: 5 Zitierungen
Artikel 7: 4 Zitierungen
Artikel 8: 3 Zitierungen
Artikel 9: 0 Zitierungen
Artikel 10: 0 Zitierungen
Artikel 11: 0 Zitierungen
Artikel 12: 0 Zitierungen

Der h-Index bezeichnet die Anzahl h der Veröffentlichungen mit jeweils mindestens h Zitierungen

Ein h-Index von 5 besagt, dass der/die Autor/in über fünf Veröffentlichungen verfügt, die jeweils mindestens fünf Mal zitiert worden sind. Im o.g. Beispiel werden fünf Artikel zumindest fünf Mal zitiert. Sechs Artikel werden jedoch nicht mindestens sechs Mal zitiert, da Artikel Nr. 6 nur über 5 Zitierungen verfügt. D.h. der h-Index beträgt 5. Artikel 1, der die meisten Zitierungen aufweist, und damit ggf. auch die höchste wissenschaftliche Bedeutung aufweist, bleibt beim h-Index unberücksichtigt. So wird sichergestellt, dass ein einzelner Artikel den Gesamtindex nicht unverhältnismäßig beeinflusst. Die Bedeutsamkeit von Artikel 1 findet jedoch in abgemilderter Weise im g-Index seine Berücksichtigung.

Der g-Index bezeichnet die Anzahl g der Veröffentlichungen mit durchschnittlich g Zitierungen

Ein g-Index von 12 besagt, dass der/die Autor/in über zwölf Veröffentlichungen verfügt, die in Summe mindestens 144 Mal zitiert worden sind. Im obigen Beispiel beträgt die Summe der Zitierungen 148. Das heißt, für 12 Artikel beträgt der Durchschnitt der Zitierungen mindestens 12 und der g-Index beträgt somit 12. In diesen Index fließt also der häufig zitierte Artikel 1 auf eine positive Weise ein, indem er geringe Zitationsraten anderer Artikel nivelliert.

Als Datengrundlage zur Berechnung dieser Faktoren sind verschiedenste Datenquellen denkbar. Hirsch selber schlägt für den h-Index das Web of Science vor.

1.3 Grundsätzliche Probleme bei Zitierindizes

Diese gängige Praxis, die Anzahl der Zitationen eines wissenschaftlichen Beitrags als Qualitätskriterium für die wissenschaftliche Leistung einer Wissenschaftlerin oder eines

Wissenschaftlers heranzuziehen, ist insofern nicht unproblematisch, als rein statistische Kriterien den Ausschlag geben. Zitationen geben einen Hinweis auf die "Popularität" unter Wissenschaftlerinnen und Wissenschaftlern - die unterschiedlich begründet sein kann - und nicht notwendigerweise auf die wissenschaftliche Relevanz einer Publikation. Der durch die Anzahl der Zitationen in wissenschaftlichen Aufsätzen gemessene Impact einer Arbeit vernachlässigt darüber hinaus den Einfluss, den eine Arbeit möglicherweise in anderen Bereichen hat (zum Beispiel in der Industrie). Weiters kann es zur Bildung so genannter "Zitierzirkel" bzw. zu häufigen Eigenzitationen kommen, die das Ergebnis verfälschen. Bei den meistzitierten Arbeiten handelt es sich auch häufig um Ausnahmen, die vor allem deshalb zitiert werden, weil es "üblich" ist, diese Arbeiten zu zitieren. Andere Arbeiten mit ebenso großem Einfluss werden dagegen nicht mehr explizit zitiert, weil ihre Inhalte bereits selbstverständlich geworden sind.

2. Datenbanken

Um die oben genannten bibliometrischen Maße eruiieren zu können, müssen Zitationsdatenbanken - i.d.R. Scopus (<http://www.scopus.com>) oder das weiter verbreitete und oben bereits erwähnte Web of Science verwendet werden. Seit 2004 bietet Google den wissenschaftlichen Suchdienst Google Scholar an, der auch die Zahl der Zitierungen jedes Eintrags angibt. Mittels "Publish or Perish", einer lokal zu installierenden Client Software, können die Ergebnisse einer Scholar Abfrage zur Berechnung verschiedenster Zitationsindizes (Publikationszahl, Zitationszahl, Zitationen pro Artikel, Zitationen pro Autor/in, Artikel pro Autor/in, h-Index, g-Index, contemporary h-index, individual h-index, age-weighted citation rate, Verteilung der Anzahl der Autorinnen bzw. Autoren pro Artikel) herangezogen werden. Dies hat gegenüber den gebräuchlichen kostenpflichtigen Datenquellen den Vorteil, dass nicht nur alleine Artikel von Fachzeitschriften berücksichtigt werden, sondern auch Monografien, Conference Proceedings oder andere wissenschaftliche Veröffentlichungen ("graue Literatur").

Die Rechercheergebnisse der Software sind jedoch naturgemäß nur so gut wie ihre Datenbasis!

Die Probleme der Nutzung des Web of Science und von Google Scholar als Datengrundlage für die Berechnung bibliometrischer Maße fasst Fell (2010) wie folgt zusammen:

Problem	Beschreibung
Unterschätzung individuellen Impacts	Publikationen außerhalb bestimmter (ISI-listed) Fachzeitschriften finden kaum Beachtung; US-/ UK-Fachzeitschriften sind überrepräsentiert. WoS-General-Search ist beschränkt auf ISI-listed Fachjournals. Über WoS-Cited-Reference sind auch von ISI-listed Fachzeitschriften zitierte nicht-ISI-listed Fachzeitschriften indiziert.
Bei nicht-ISI-listed Fachzeitschriften nur Erstautoren	Bei der Zählung von Zitationen (WoS-Cited-Reference) werden nur Erstautorinnen bus. -autoren berücksichtigt, was bei Koautoren bzw. -innen für Impact-Unterschätzung sorgt.
Mangelhafte Zitationsaggregation	Es gelingt oft nicht, Zitationen, die sich durch minimale Schreibfehler unterscheiden, zu aggregieren und dem richtigen Werk zuzuordnen.

Tabelle 1: Probleme beim Web of Science (WoS)

Problem	Beschreibung
Nichtwissenschaftliche Zitationen	Es werden auch nichtwissenschaftliche Zitationen (z. B. in Hausarbeiten) gezählt.
Unklare, ungleiche Abdeckung	Google gibt kaum etwas über die Abdeckung preis; vollständig ist sie jedoch bei weitem nicht. Bestimmte Fächer profitieren (z. B. Psychologie), andere sind benachteiligt (z. B. Chemie). Auch jüngere Publikationen (≥ 1990) profitieren.
Unsinnige Ergebnisse	Der Parser erzeugt gelegentlich sinnlose Ergebnisse (z. B. unsinnige Namen von Autoren bzw. Autorinnen).
Publikationstypen	Eine Unterscheidung zwischen verschiedenen Publikationstypen ist nicht/kaum (s. u.) möglich.
Inkonsistente Zitationen	Zitationen sind teilweise inkonsistent, z. B. hinsichtlich Schreibweise oder Vollständigkeit.
Mangelnde Verschlagwortung, Klassifikation	Suchmodi: Stichwortsuche (irgendwo, im Titel) Name des Autors/der Autorin, Publikationsname, Veröffentlichungsjahr, Domain, URL, Dateityp
Wenig komplexe Suchsyntax	Suche nach Stoppwörtern (+), Ausschluss bestimmter Begriffe (-), Boolesches Verodern (OR,), Phrasensuche („“)

Tabelle 2: Probleme bei Google Scholar

Weitere Probleme ergaben sich bei einer am Department für Interaktive Medien und Bildungstechnologien durchgeführten beispielhaften Recherche¹ mittels der Software "Publish or Perish" wie folgt:

- Anzeigen oft doppelter Werke

Hier trat das Problem auf, dass die doppelt vorkommenden Werke nicht einfach deselektiert werden können. Dies ist nur möglich, wenn die Werke nicht zitiert wurden.

- Verschiedene Publikationstitel

Hier trat das Problem auf, dass Werke sowohl in Deutsch als auch in Englisch publiziert werden.

- Abweichungen des Suchergebnisses von der tatsächlichen Publikationsliste

Bei *allen* Wissenschaftlerinnen bzw. Wissenschaftlern wich das Ergebnis der tatsächlichen Publikationsliste von dem Ergebnis der Suche von „Publish or Perish“ ab. Es wurden meist deutlich weniger Werke gefunden, als der/die Autor/in laut Publikationsliste veröffentlicht hat.

3. Schlussfolgerungen

Neben den in den obigen Ausführungen dargestellten allgemeinen Problemen bzw. Limitierungen von bibliografischen Maßen zur Evaluierung der Publikationsleistung von

¹ Die Recherche wurde von Florian Eder exemplarisch mit sieben Wissenschaftlerinnen bzw. Wissenschaftlern aus dem Bereich der Pädagogik und des technologieunterstützten Lernens durchgeführt.

Wissenschaftlerinnen und Wissenschaftlern, ist insbesondere auch die Nutzung von Google Scholar als Datenbasis mit größter Vorsicht zu betrachten. Allerdings kann mittels der vorgestellten Software - sofern ein sorgfältiger Abgleich mit der Publikationsliste des betreffenden Autors bzw. der betreffenden Autorin erfolgt - eine einfache und kostengünstige Analyse vorgenommen werden, deren Datenbasis über die alleinige Nutzung von Zeitschriftenartikeln hinausgeht. h- und g-Index beispielsweise können so im direkten Vergleich mehrerer Wissenschaftler/innen Anhaltspunkte für die Publikationsleistung bieten.

Festzuhalten bleibt, dass bibliografische Maße nur einen Indikator darstellen können, um die wissenschaftliche Qualität der Publikationen von Autorinnen bzw. Autoren zu messen, und eine gutachterliche Tätigkeit bzw. eine qualitative Evaluierung nicht zu ersetzen in der Lage sind.

4. Verwendete Quellen:

Fell, C. (2010). Publish or perish und Google Scholar – ein Segen? Abgerufen unter http://www.zpid.de/pub/research/2010_Fell_Publish-or-Perish.pdf [20.09.2010].

<http://de.wikipedia.org/wiki/Zitationsanalyse> [20.09.2010].

<http://de.wikipedia.org/wiki/H-Index> [20.09.2010].

http://de.wikipedia.org/wiki/Impact_Factor [20.09.2010].

<http://wokino.com/> [20.09.2010].

http://de.wikipedia.org/wiki/Web_of_Science [20.09.2010].

<http://www.harzing.com/pop.htm> [20.09.2010].

Noruzi, A. (2005). Google Scholar: The New Generation of Citation Indexes. *International Journal of Libraries and Information Services*, 55(4), 170-180. Abgerufen unter: <http://www.librijournal.org/pdf/2005-4pp170-180.pdf> [20.09.2010].