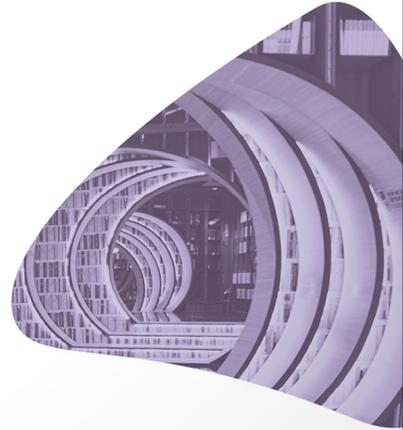




Melusina Press
Humanities



Reputation ohne Paywall?

Wissenschaftliches Publizieren im digitalen Wandel

*Yuliya Fadeeva, Simone Franz, Torsten Kahlert, Dario Kampkaspar,
Melanie Seltmann, Timo Steyer, Niels-Oliver Walkowski (Hrsg.)*

Messen, zählen, wiegen
Vom Journal Impact Factor zu den Responsible Metrics in den
Naturwissenschaften

Birgid Schindwein

Reputation ohne Paywall? Wissenschaftliches Publizieren im digitalen Wandel
Herausgegeben von Yuliya Fadeeva, Simone Franz, Dario Kampkaspar,
Melanie Seltmann, Timo Steyer und Niels-Oliver Walkowski



Melusina Press, 2025

Veröffentlicht durch die Universität Luxemburg - **Melusina Press**, 2025
11, Porte des Sciences
L-4366 Esch-sur-Alzette
<https://www.melusinapress.lu>

Verlagsleitung: Niels-Oliver Walkowski, Johannes Pause
Lektorat: Carolyn Knaup, Niels-Oliver Walkowski
Cover und Layout: Valentin Henning, Erik Seitz
Umschlagsbild: @vnwayne_fan; Foto 2020 veröffentlicht auf Unsplash

Die digitale Version dieser Publikation steht unter <https://www.melusinapress.lu> frei zur Verfügung.
Das PDF und die Druckvorlage werden mit Hilfe von princeXML erzeugt.

Bibliografische Information der Nationalbibliothek Luxemburg: Die Nationalbibliothek Luxemburg verzeichnet diese Publikation in der Luxemburgischen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über bnl.public.lu abrufbar.

ISBN (Online)	978-2-919815-83-8
ISBN (PDF)	978-2-919815-84-5
ISBN (EPub)	978-2-919815-85-2
DOI (Publikation)	10.26298/1981-5838-mzwx
DOI (Band)	10.26298/1981-5838

Das vorliegende Werk steht unter einer CC BY-SA 4.0 Lizenz. Informationen zu dieser Lizenz finden Sie unter <https://creativecommons.org/licenses/by-sa/4.0/deed.de>. Die in diesem Werk enthaltenen Bilder und Ressourcen unterliegen der selben Lizenz, sofern sie keiner anderen Quelle entnommen wurden oder mit einer anderen Lizenz versehen sind.



Messen, zählen, wiegen – Vom Journal Impact Factor zu den Responsible Metrics in den Naturwissenschaften

Birgid Schindwein

1. Einführung

„Man muss messen, was messbar ist, und messbar machen, was noch nicht messbar ist.“
Galileo Galilei zugeschrieben.

„Das ist ja das Schöne an den Naturwissenschaften, da braucht man keine Meinung, wir brauchen Messungen.“
Harald Lesch (2024)

In den Naturwissenschaften wird gezählt, gemessen und gewogen. Den Wissenschaftler:innen sind quantitative Methoden so vertraut, dass es ihnen ganz selbstverständlich ist, ihre eigenen wissenschaftlichen Leistungen zu messen oder messen zu lassen. Viele der heute häufig genutzten Methoden der Forschungsbewertung stammen aus der Bibliometrie. Dabei entstand die Bibliometrie ursprünglich in der Bibliothekswissenschaft. Man wollte die Nutzung bibliothekarischer Materialien und Dienstleistungen mit mathematisch-statistischen Methoden analysieren, um beschränkte finanzielle Mittel effizienter einsetzen und neue Angebote machen zu können (Reitz 2021).

Als vermessende Lehre arbeitet die Bibliometrie quantitativ, sie *bewertet* also im Sinne des Qualitätsmanagements zum Beispiel die Anzahl der Publikationen einer Institution binnen eines Jahres. Der quantitative Beitrag der Bibliometrie zur Forschungsbewertung mittels Indikatoren ergänzt definitionsgemäß und in der Theorie die qualitative Bewertung zum Beispiel mittels Peer Review.

Zwei Annahmen werden in der Bibliometrie zu Grunde gelegt:

1. *Publizieren* ist das zentrale Element wissenschaftlicher Arbeit.
2. *Zitationen* zeigen, wie wissenschaftliche Arbeit wahrgenommen wird.

Aus beiden Größen, der Anzahl an Publikationen und der Zahl an Zitationen für diese Publikationen, leitet die bibliometrische Theorie den wissenschaftlichen Impact eines Publikationsorgans, einer Institution, Arbeitsgruppe, Einzelperson oder einer Publikation ab. Der englische Begriff *impact* darf in diesem Zusammenhang nicht mit *influence* gleichgesetzt werden. Eine Zitation repräsentiert die Wirkung, die eine Publikation erzeugt hat, nicht den Einfluss, den sie ausgeübt hat oder gar in der Zukunft ausüben wird. Reputation wird in der bibliometrischen Theorie durch Zitationen nicht erzeugt.

Garfield entwickelte den Journal Impact Factor (JIF) 1955 als Metrik für den Einfluss von (zunächst nur) naturwissenschaftlichen und medizinischen Zeitschriften. 2005, also ein halbes Jahrhundert später, führte der Physiker Hirsch den nach ihm benannten h-Index für Personen ein. Dies war eine gezielte Aktion gegen den Einsatz des JIF, also einer Zeitschriftenmetrik, für Personen, der sich ab den 1970er Jahren schnell eingebürgert hatte. Auch bei der unangebrachten Verwendung des JIF für Personen hatten die Naturwissenschaften und die Medizin eine Vorreiterrolle übernommen. Mit den Journal Citation Reports, in denen jährlich das JIF-Ranking der Zeitschriften publiziert wird, steht das ideale Instrument – hinter der Bezahlwand – zur Verfügung. Und hier kommt auch das erste Mal *Reputation* ins Spiel, zunächst die von Zeitschriften. Erleichtert wurde und wird die Fokussierung auf eine Zeitschriftenmetrik, da die wissenschaftliche Kommunikation in den Naturwissenschaften traditionsgemäß hauptsächlich über Zeitschriftenaufsätze erfolgt (Haustein 2010). In Ermangelung einer Metrik für Qualität wurde und wird der JIF aber auch bereitwillig und unkritisch als einfach verständliches und einfach zu ermittelndes „Surrogat“ (Baethge 2012) herangezogen – und das nicht nur für Zeitschriften, sondern auch für alle anderen Entitäten wie für Personen oder Einzelartikel. Obwohl durch die bibliometrische Forschung versucht wurde, den groben Vereinfachungen und falschen Anwendungen mit Korrekturen, z. B. mit Fachgewichtungen, oder durch andere Methoden wie Netzwerkanalysen oder Visualisierungen zu begegnen, sind der JIF und der h-Index bis heute die bekanntesten und am häufigsten verwendeten Metriken in der Forschungsbewertung – nicht nur in den Naturwissenschaften.

2. Voraussetzungen – Datengrundlagen für die Erhebung bibliometrischer Kennzahlen

Für bibliometrische Analysen liefern Zitationsdatenbanken – Literaturdatenbanken, die zusätzlich die zitierte Literatur der referenzierten Publikationen enthalten – die Datengrundlage. Die Aktualität und die Größe der Datenbanken sowie vor allem die Korrektheit aller enthaltenen Daten sind Grundbedingungen für eine valide Aussage der damit erstellten bibliometrischen Analysen. Zum gegenwärtigen Zeitpunkt werden in den meisten Zitationsanalysen die Datenbanken Google Scholar, Scopus und Web of Science (WoS) genutzt. Google Scholar steht frei im Internet zur Verfügung, die beiden anderen Datenbanken müssen lizenziert werden. Die ebenfalls frei verfügbare Datenbank OpenAlex wurde erst 2022 als Literaturdatenbank und für bibliometrische Analysen gestartet, sie wird bisher erst vereinzelt für letztere eingesetzt. Ausnahme ist das CWTS Leiden Ranking Open Edition, das seit 2024 die wissenschaftliche Performanz von über 1.500 Universitäten aus aller Welt basierend auf Daten aus OpenAlex ermittelt (Waltman et al. 2024).

Im Folgenden soll untersucht werden, ob das Vertrauen in die Daten aller dieser Zitationsdatenbanken und die Methoden angebracht ist oder ob die Verwendung der einfachen Metriken, besonders des Journal Impact Factors aber auch des h-Indexes, in Frage gestellt werden muss.

2.1 Datenqualität

Bei Google Scholar erzeugt die Erstellung der Datenbank durch automatisierte Webcrawler und fehlende Qualitätskontrolle immer wieder Inkonsistenzen, Fehler und Dubletten (Harzing and Alakangas 2016), die nicht behoben werden. Autorenprofile müssen von Wissenschaftler:innen selbst angelegt werden. Sie sind auch dafür verantwortlich, dass ihnen ihre Publikationen und nur ihre Publikationen zugerechnet werden. Die automatische Erkennung und Zuordnung neu indexierter Publikationen zum eigenen Profil sollte abgeschaltet werden, leider ist sie Standard. Man kann bei Publikationen, die in einem Verlag erschienen sind, davon ausgehen, dass die Metadaten in der PDF-Datei vorhanden und korrekt sind, und damit auch korrekt erkannt werden, doch ist das bei Preprints nicht gewährleistet. Bei Namensgleichheit wird es von Google Scholar den betroffenen Autor:innen überlassen, ihre Profile korrekt und aktuell zu halten. Sowohl die Qualität der Publikationen als auch die Datenqualität ist in Google Scholar sehr verschieden. Dennoch wird die Datenbank gerne vor allem zur Erhebung des h-Indexes für Einzelpersonen verwendet, da sie durch ihre schiere Größe die höchsten Werte generiert.

OpenAlex ist ebenfalls nicht selektiv beim Indexieren von Publikationen, so finden sich zum Beispiel Preprints (OurResearch 2024) aber auch Artikel aus nicht vertrauenswürdigen Zeitschriften in der Datenbank. Die Hauptdatenquelle für OpenAlex ist Crossref, eine der Registrierungsagenturen für Digital Object Identifiers (DOIs). Verlage liefern die Metadaten von Publikationen einschließlich der Literaturlisten automatisch an Crossref. Dies ist nicht nur ein Einfallstor für Publikationen zweifelhaften Inhalts (Richtig et al. 2023), sondern wird auch zum Einschleusen zusätzlicher Zitationen, die überhaupt nicht im Artikel vorkommen, benutzt (Besançon et al. 2024). Bei der Disambiguierung der Autoren stützt sich OpenAlex zum Beispiel auf ORCID. Sind in den gelieferten Metadaten keine angegeben, ist der Algorithmus zur Erzeugung von Autorprofilen sehr großzügig, so dass sich Autoren mit häufigen Namen dann in einem Profil wiederfinden (Culbert et al. 2024). In einer Bewertung der Eignung von OpenAlex für bibliometrische Analysen, die in Zusammenarbeit mit OurResearch, dem Produzenten von OpenAlex, erstellt wurde, mahnen Alperin et. al. (2024) noch zur Vorsicht. So ist OpenAlex immer noch Untersuchungsobjekt von Bibliometriker:innen, um den Umfang und die Datenqualität zu testen und mit den traditionellen Datenbanken zu vergleichen. Der Umfang von OpenAlex übertrifft sowohl Scopus als auch WoS. Die Aufnahme von Publikationen aus Raubzeitschriften kann zwar durch Filterung auf im Directory of Open Access Journals (DOAJ) enthaltene Publikationsorgane kompensiert werden, in die Zitationsdaten gehen sie aber dennoch ein. Auch die eindeutige Erkennung und Disambiguierung bei Personen und Institutionen ist – Stand Juni 2025 – immer noch nicht so gut, dass valide bibliometrische Analysen möglich wären. Das Centrum voor Wetenschaps- en Technologiestedies (CWTS) in Leiden, NL nutzt erhebliche eigene Ressourcen und auch weitere Datenquellen für die korrekte Zuordnung von Publikationen zu den untersuchten Universitäten (CWTS 2024).

Die Firmen Clarivate und Elsevier bewerben ihre Datenbanken WoS beziehungsweise Scopus mit hohen Qualitätsansprüchen an die Selektivität und Sorgfalt der Indexierungen. Die beiden Datenbanken sollten demzufolge wesentlich weniger Fehler enthalten

als Google Scholar und OpenAlex. Leider sind jedoch die bekannten Schwächen bei der Disambiguierung von Zeitschriften, Institutionen und Personen im WoS (Rehs 2021) auch nach Jahren nicht behoben. Die Algorithmen zur Erstellung von Personenprofilen sind immer noch so schwach, dass die generierten Profile für eine bibliometrische Auswertung nur mit großer Vorsicht und profunder Kenntnis der Sachverhalte zu nutzen sind. Es werden mehrere Profile pro Autor:in angelegt oder es werden die Publikationen mehrerer Autor:innen in ein Profil zusammengelegt. Sogar Kurationen der Autor:innen selbst durch das Claimen und Prüfen des eigenen Profils werden immer wieder überschrieben. Werden ohne Umsicht, Sachkenntnis und Kontrolle einfach die Autorenprofile des WoS zur Analyse herangezogen, sind die ermittelten Metriken in vielen Fällen wertlos in der Aussage und ungerecht für die Person, zu der sie erhoben wurden. Das gleiche gilt für die korrekte Zuordnung von Publikationen zur Forschungsstätte. Auch hier sind zwar Korrekturen durch die Institutionen selbst möglich, indem alle kursierenden Namensvarianten der eigenen zugeordnet werden können. Es ist jedoch nicht transparent, welche Institutionen diese Möglichkeit nutzen.

The screenshot shows the Scopus search results interface. At the top, it displays the Scopus logo and the text "1,587 documents have cited:". Below this, the search criteria are shown: "Dependency and development in Latin America" by "Cardoso F.H., Faletto E." with a red circle around the year "(2024)". The citation text is "Dependency and Development in Latin America, , pp. 1-254.". The interface includes a search bar, a "Refine results" section with "Limit to" and "Exclude" buttons, and a "Document type" filter. A "Year" filter is expanded, showing a list of years from 2018 to 2025 with corresponding document counts: 2025 (3), 2024 (84), 2023 (83), 2022 (61), 2021 (63), 2020 (62), 2019 (90), and 2018 (80). On the right side, there is an "Analyze search" section with a "Docume" filter and two document types: "1 Decolon late neo" and "2 Social, c patrimo governa".

Abbildung 1: Screenshot aus der Datenbank Scopus vom 27. 09.2024

Bekanntermaßen wurden mehrfach Raubzeitschriften mit ihren Artikeln von zweifelhafter Qualität in die Datenbank Scopus aufgenommen (Cortegiani et al. 2019). Durch eine Verschärfung der Aufnahmekriterien und deren regelmäßig Überprüfung wurde die weitere Indexierung von Artikeln aus diesen Publikationsorganen gestoppt, aber nicht rückgängig gemacht. So sind die einmal indexierten Artikel nach wie vorhanden und gehen auch in die Daten für Zitationsanalysen ein.

Auch in Scopus kann man mit einfachen Suchen systematische Fehler in Meta- und Zitationsdaten finden. Wenn man Ende September 2024 in Scopus Dokumente mit dem Publikationsjahr 2024 suchte, fiel auf, dass unter den Top-Treffern mehrere Bücher und Buchkapitel waren. Dies wäre sehr erstaunlich, da Zeitschriftenartikel in der Regel wesentlich schneller und häufiger zitiert werden als Bücher und Buchkapitel.

Das falsche Publikationsjahr für den Titel *Dependency and development in Latin America* von Cardoso und Faletto (laut Scopus von 2024, korrektes Publikationsjahr der neuesten Auflage 1979) führt zu der skurrilen Situation, dass der Titel laut Zitationsdatenbank schon Jahrzehnte vor seinem angeblichen Erscheinen zitiert wurde (s. Abb. 1). Selbst wenn man den Nachdruck der 8. Auflage von 2020 heranzieht, der 2021 auch als E-Book-Version veröffentlicht wurde, sind die bibliographischen und die Zitationsangaben falsch. Keinem Bachelorstudierenden würde man ein Zitat mit solch falschen bibliographischen Angaben durchgehen lassen. Die Monographie wird in Scopus auch im Juni 2025 noch immer als eine der meistzitierten Publikationen des Erscheinungsjahrs 2024 gelistet. Leider ist sie kein Einzelfall (siehe auch Abbildung 4). Eine Datenkorrektur ist bis zur Fertigstellung des Manuskripts Ende Juni 2025 trotz Meldung des Fehlers nicht erfolgt.

Mit einer einfachen Plausibilitätsprüfung wären solche Fehler zu vermeiden. Auch dass ein Artikel sich angeblich selbst zitiert, sollte mittels Plausibilitätstest ermittelt und die Aufnahme solcher falschen Zitationsdaten vermieden werden können. Im folgenden Beispiel sieht man die Zirkelzitation:

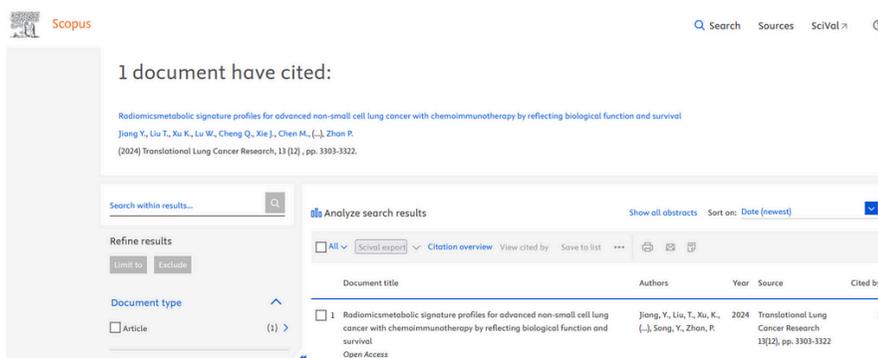


Abbildung 2: Screenshot aus der Datenbank Scopus vom 27.09.2024

Ein besonderes Augenmerk sollte außerdem auf der Prüfung von Bewertungen wie *highly cited* liegen. Beispielsweise bleibt im Vergleich mit den Zitationszahlen auf den

Plattformen Google Scholar und Scopus rätselhaft, wie es im WoS dazu kommen kann, dass der Artikel *Search for sources of high-energy neutrons with four years of data from the IceTop Detector* (Aartsen et al. 2016) zum *Highly Cited Article* geadelt wird, ohne dass die Zitationsdaten verifiziert werden.

Zitationszahlen des selben Artikels in Google Scholar, Scopus und WoS

The image contains three screenshots of search results for the article "SEARCH for SOURCES of HIGH-ENERGY NEUTRONS with FOUR YEARS OF DATA from the ICETOP DETECTOR".

- Google Scholar (top):** Shows a search result with the citation count "Zitiert von: 10" circled in red.
- Scopus (middle):** Shows a table with columns: Document title, Authors, Source, Year, and Citations. The citation count "6" is circled in red.
- WoS (bottom):** Shows a search result with the citation count "792 Citations" circled in red.

Abbildung 3: Screenshots aus den Datenbanken Google Scholar, Scopus und Web of Science vom 27.09.2024

Nun könnte man argumentieren, dass die genannten Fehler zwar ärgerlich, aber nicht ausschlaggebend für die ermittelten Metriken sind, da sie nur einen sehr geringen Anteil der in den Datenbanken enthaltenen Publikationen betreffen. Autor:innen, deren Profile nicht korrekt sind, und die eine Karrierechance aufgrund unvollständiger Daten verpasst haben, sind doppelte Opfer: die der Datenbankunzulänglichkeiten und eventuell missbräuchlicher Verwendung von Metriken. Nicht nur der h-Index einer Autorin kann betroffen sein, sondern sogar die fachnormalisierte Metrik einer Einzelpublikation eines Autors. Für das Publikationsjahr 2024 sind in Abbildung 4 die in Scopus enthaltenen Bücher aus den Geisteswissenschaften mit den höchsten Zitationen gelistet. Alle abgebildeten Publikationen sind wesentlich früher erschienen und liegen zum Teil schon in mehreren Auflagen vor, keine ist jedoch 2024 erschienen. Allein diese wenigen Publikationen treiben die durchschnittliche Anzahl von Zitationen in die Höhe: Geisteswissenschaftliche Bücher mit (angeblichem) Publikationsjahr 2024 sollen bisher im Durchschnitt 2,3-mal zitiert worden sein, wogegen geisteswissenschaftliche Zeitschriftenartikel aus 2024 im Schnitt seit ihrem Erscheinen nur 0,8-mal zitiert wurden. Der hohe Durchschnitt geht zu Lasten von Büchern, die wirklich 2024 erschienen sind, und benachteiligen damit natürlich auch deren Autor:innen.

Publications in the World

Within: Arts and Humanities | Year: 2024 | Applied filters: Book × Reset filter Export

9,803 publications | Save as Publication Set

	Title	Authors	Year	Scopus Source	Citations
<input type="checkbox"/> Morris, P. 18	Dependency and development in Latin America	Cardoso, F.H., Faletto, E.	2024	Dependency and Development in Latin America	1,502
<input type="checkbox"/> Basker, J.G. 15					
<input type="checkbox"/> Deguchi, Y. 15					
<input type="checkbox"/> Owens, W.R. 15	Gut Reactions: A Perceptual Theory of Emotion	Prinz, J.J.	2024	Gut Reactions:A Perceptual Theory of Emotion	1,028
<input type="checkbox"/> Furbank, P.N. 13					
Show more					
View all					
Institutions					
<input type="checkbox"/> University of Oxford 115	The Eyes of the Skin: Architecture and the Senses	Pallasmaa, J.	2024	The Eyes of the Skin: Architecture and the Senses	1,002
<input type="checkbox"/> University of Cambridge 75	The Great Escape: Health, Wealth, and the Origins of Inequality	Deaton, A.	2024	The Great Escape: Health, Wealth, and the Origins of Inequality	910
<input type="checkbox"/> University of Edinburgh 66	Mind A Brief Introduction	Searle, J.R.	2024	Mind a Brief Introduction	538
<input type="checkbox"/> City 54	Philosophy of Mathematics and Natural Science	Weyl, H.	2024	Philosophy of Mathematics and Natural Science	450
	Pure theory of law	Kelsen, H.	2024	Pure Theory of Law	418

Abbildung 4: Screenshot aus der Datenbank Scopus vom 26.06.2025

Dass bibliographische und Zitationsdatenbanken nicht fehlerfrei sind, ist bekannt (Zhu, Hu, and Liu 2019; Liu 2021; Donner 2024). Alle hier beschriebenen Fehler sind mit ganz einfachen Suchanfragen zu ermitteln. Das Ausmaß, in dem Fehler bei Meta- und Zitationsdaten vorkommen, überrascht aber immer wieder und auch, dass offensichtlich selbst einfache Gegenmaßnahmen wie die oben geforderten Plausibilitätsprüfungen nicht ergriffen werden. Falsche Meta- und Zitationsdaten können nie zu korrekten bibliometrischen Analysen führen.

2.2 Statistische Validität

Der JIF zieht schon lange eine erstaunliche Liste von Kritiken (Seglen 1997) auf sich. Der massivste Vorwurf dieser Liste ist, dass der JIF als arithmetischer Mittelwert keine statistisch valide Aussage über die einzelnen Werte macht, da die Voraussetzung der Normalverteilung der Daten verletzt wird. Auch in Zeitschriften mit hohen JIFs werden bis zu 25% der Dokumente nie zitiert. Zitationszahlen sind sehr schief verteilt, deren Mittelwert ist damit von statistisch unqualifizierter Beliebigkeit. In den Journal Citation Reports von Clarivate wird deshalb in der Verteilungsgraphik der Zitationen der Zeitschriften korrekterweise auch der Median verwendet und nicht der Mittelwert. Die Ablehnung als valide Metrik gilt in gleicher Weise für den CiteScore, die Zeitschriftenmetrik von Scopus. Auch der CiteScore ist ein arithmetisches Mittel und damit nicht geeignet, Zitationsdaten einer Zeitschrift statistisch korrekt zu beschreiben.

Generell darf das arithmetische Mittel bei Zitationsdaten nicht verwendet werden, außer sie sind normalverteilt. Untersuchungen in mehreren Fächern haben gezeigt, dass Zitationen lognormal-verteilt sind (Stringer et al. 2010)

Der h-Index bringt die Anzahl der Publikationen eines Autors mit den Zitationen für diese Publikationen in einen recht willkürlichen Zusammenhang. Warum benötigt man für einen h-Index von 8 genau 8 Publikationen mit mindestens 8 Zitationen? Könnte man nicht auch 4, 16 oder gar 80 Zitationen fordern? Es gibt keine Notwendigkeit oder Begründung für die geforderte Zitationszahl.

2.3 Automation Bias

Der Automation Bias bewirkt, dass die einfache Verfügbarkeit einer automatischen Entscheidungshilfe zu deren immer häufigerem Einsatz führt. So ist der h-Index in den gängigen Zitationsdatenbanken jederzeit entweder direkt verfügbar oder sehr leicht zu ermitteln. Entsprechend beliebt ist die Metrik sowohl bei den (Natur-)Wissenschaftler:innen und Mediziner:innen selbst, als auch in der Forschungsbewertung. Schon bald nach seiner Einführung im Jahr 2005 und bis heute reißt jedoch die Kritik auch am h-Index nicht ab (Nielsen 2008; Waltman and van Eck 2012; Yong 2014; Koltun and Hafner 2021; Haustein 2023). Um diese Schwächen auszugleichen, wurden von Bibliometriker:innen (Egge 2007; Bornmann, Mutz, and Daniel 2008) fast ebenso viele Varianten des h-Indexes entwickelt. Verwendet wird keine davon.

In der Datenbank Web of Science wird in der Einzeltrefferanzeige jedes Zeitschriftendokuments auch der JIF der Zeitschrift angezeigt, in der die Publikation erschienen ist – ready to use. Dabei sagt der JIF überhaupt nichts über den Einzelartikel aus, vor allem nichts über dessen Qualität.

Wenn die Anzahl der zu erledigenden Gutachten erdrückend hoch ist, ist auch die Verlockung hoch, die in den Datenbanken angebotenen Zahlen für eine Beurteilung heranzuziehen. Ein Kreislauf aus Nutzung und Akzeptanz entsteht.

3. Bewertung von Forschungsleistungen

All diese Unzulänglichkeiten sind längst bekannt und anerkannt. Und doch wird immer noch in so mancher Habilitationsordnung, vor allem in der Medizin, die Publikation einer Mindestanzahl von Artikeln in Zeitschriften mit einem JIF über einem bestimmten Schwellenwert gefordert. Aus einem hohen JIF wird die Reputation einer Zeitschrift abgeleitet und aus dieser angenommenen Reputation der Zeitschrift die Qualität des in ihr veröffentlichten Artikels und der Forschung einer Person.

In Bayern ist die Leistungsorientierte Mittelvergabe (LOM) an den Universitätskliniken u. a. abhängig vom JIF der Zeitschriften, in denen publiziert wurde. Bei der Erfassung der Publikationen werden die JIFs der Zeitschriften automatisch eingespielt. Wenn eine Zeitschrift nicht im WoS gelistet ist, hat sie keinen JIF. Das macht die Publikation für dieses Verfahren wertlos. Auch Kongressbeiträge scheiden aus, wenn die Procee-

dingsbände keinen Zeitschriftenstatus haben, selbst wenn die Konferenz im Teil *Conference Proceedings Citation Index* des WoS referenziert ist.

Es gibt wohl kein universitäres Berufungs-, Preisverleihungs- oder Aufstiegsverfahren (in den Naturwissenschaften) ohne zumindest eine Spalte für den h-Index in der Übersichtstabelle über die Kandidat:innen. Immer wieder finden sich in Kommissionen Tischvorlagen wie die folgende:

Tabelle 1: Beispiel einer Tischvorlage

AUTORIN	ANZAHL PUBLIKATIONEN	ANZAHL ZITATIONEN	H-INDEX
A	282	72.002	51
B	284	3.374	29
C	369	18.455	57
D	150	1.859	24

Welche Aussage kann man der Tabelle entnehmen? Als einziges kann man sagen, dass alle Autorinnen schon viel publiziert haben. Da aber sowohl die Quelle der Daten als auch das Erhebungsdatum fehlen, ganz zu schweigen von den Fachgebieten oder dem wissenschaftlichen Alter der Autorinnen, verbietet sich jede Bewertung mit Hilfe des h-Indexes, die karriererelevant sein könnte.

4. Herausforderungen und Grenzen

Bei vielen und vor allem bei vielen jungen Wissenschaftler:innen werden mit der Reduktion ihrer Forscherpersönlichkeit und Forschungsleistung auf eine Zahl ethische Grenzen überschritten. Letztere sind zu jung, um schon viele Arbeiten publiziert zu haben. Ihre Publikationen sind zu jung, um schon viele Zitationen gesammelt zu haben. So kommt es bei Bewertungen am Karriereanfang zum reflexartigen Rückgriff auf die Anzahl der Publikationen in *hochkarätigen* Zeitschriften, auch wenn die Deutsche Forschungsgemeinschaft 2019 in Leitlinie 15 ausdrücklich feststellt: „Die wissenschaftliche Qualität eines Beitrags hängt nicht von dem Publikationsorgan ab, in dem er öffentlich zugänglich gemacht wird.“

Bereits 2008 hielten mehrere mathematische Vereinigungen fest: „The sole reliance on citation data provides at best an incomplete and often shallow understanding of research—an understanding that is valid only when reinforced by other judgments. *Numbers are not inherently superior to sound judgments.*“ (Adler, Ewing, and Taylor 2008) Die deutliche Kritik an der *culture of numbers* und die Forderung, dass die Bewertung von Forschungsleistung um Methoden ergänzt werden müsse, die über reine Zahlen hin-

ausgehen, wurde veranschaulicht mit dem Irrglauben, nur anhand des Gewichts einer Person deren Gesundheit beurteilen zu können.

Immer mehr Widerstand formierte sich gegen diese Praktiken, zunächst unter Betroffenen, bei denen der Leidensdruck am größten ist. So wurde 2012 anlässlich einer Konferenz der American Society of Cell Biologists die *San Francisco Declaration on Research Assessment* (DORA) verabschiedet. Sie richtet sich in erster Linie gegen die Verwendung von Zeitschriftenmetriken für Personen. In 18 Empfehlungen wird sehr detailliert gefordert, dass zum Beispiel der JIF nicht zur Beurteilung von Einzelpersonen herangezogen werden darf.

Das *Leiden Manifesto* (Hicks et al. 2015) stammt von Scientometiker:innen u. a. der Universität Leiden. Es fordert erneut, dass quantitative Metriken nur zur Unterstützung qualitativer Beurteilungen herangezogen werden dürfen. Weiterhin wird unter anderem verlangt, dass Indikatoren regelmäßig auf ihre Aussagekraft hin überprüft werden und gegebenenfalls ersetzt werden müssen.

Die Coalition for Advancing Research Assessment (CoARA) wurde im Januar 2022 von über 350 Institutionen aus über 40 Ländern gegründet. Am 12.06.2025 hatten bereits 844 Institutionen ihr *Agreement on Reforming Research Assessment* (Coalition for Reforming Research Assessment 2022) unterzeichnet und sich damit zu einem umfangreichen Maßnahmenpaket verpflichtet. Dies ist umso bemerkenswerter, als die Unterzeichnung mit der Bereitstellung erheblicher Ressourcen für die Umsetzung der geforderten Maßnahmen verbunden ist.

Während eine Unterschrift unter DORA nur eine Absichtserklärung und ein Bekenntnis zu den Empfehlungen ist, ist die Beteiligung bei CoARA eine Verpflichtung. So verwundert es nicht, dass bis Juni 2025 nur 78 deutsche Institutionen DORA unterschrieben haben und gar nur 44 das Agreement von CoARA.

5. Responsible Metrics

Die Verwendung von Metriken bedeutet Verantwortung: Nur wer gut informiert ist, kann bibliometrische Indikatoren korrekt einsetzen. Jede:r Wissenschaftler:in und jede:r Wissenschaftsmanager:in muss heute über die Berechnungsmethoden, die Aussagekraft und die Unzulänglichkeiten der verwendeten Metriken Bescheid wissen. Jede:r Dozent:in und jede:r Betreuer:in im wissenschaftlichen Betrieb sollte heute die Bedeutung von Metriken zur Messung von Forschungsleistungen kennen, sowie die verantwortungsvolle Anwendung praktizieren und dem wissenschaftlichen Nachwuchs vermitteln. Das setzt in manchen akademischen Kreisen immer noch Mut voraus.

In einigen Fächern spielen Metriken wie der h-Index für die Forschungsbewertung von Personen keine Rolle mehr, in anderen wie der Medizin und der Biologie werden sogar die JIFs noch herangezogen. Fächer, in denen noch nie mit diesen Metriken zur Bewertung gearbeitet wurde, müssen sich mittlerweile nicht mehr dazu treiben lassen. Stattdessen stehen fortschrittliche Methoden wie zum Beispiel die Visualisierung der Zitationsperzentile von Publikationen zur Verfügung. Mit einem Beamplot (Bornmann and Marx 2014) wird entweder für alle Publikationen einer Person oder für Publikatio-

nen aus einem ausgewählten Zeitraum dargestellt, wie wenig oder hoch sie im Vergleich zu Publikationen desselben Fachgebiets, aus demselben Jahr und vom selben Dokumenttyp zitiert werden. Diese Grafiken vermitteln einen wesentlich besseren Eindruck von der Publikationstätigkeit und dem Impact als eine einzige Zahl und machen auch Vergleiche möglich.

Bibliometrische Indikatoren geben als quantitative Indikatoren kein Qualitätsurteil ab, sie sagen noch nicht einmal etwas über den wissenschaftlichen Einfluss einer Publikation, eine:r Forscher:in oder einer Universität aus und haben nichts mit Reputation zu tun. Sie beruhen auf Zitationen und spiegeln höchstens die Sichtbarkeit wider. Bibliometrische Methoden können als *eine* Facette zur Forschungsbewertung beitragen. Bibliometrie ist dafür da, um zu erkennen, wo man Fragen stellen muss.

Literatur

- Aartsen, M. G., K. Abraham, M. Ackermann, J. Adams, J. A. Aguilar, M. Ahlers, M. Ahrens et al. 2016. „Search for Sources of High-Energy Neutrons with Four Years of Data from the IceTop Detector.“ *ApJ* 830 (2): 129. <https://doi.org/10.3847/0004-637X/830/2/129>.
- Adler, Robert, John Ewing, and Peter Taylor. 2008. „Citation Statistics: A Report from the International Mathematical Union (IMU) In Cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) And the Institute of Mathematical Statistics (IMS).“ Accessed January 26, 2025. <http://www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf>.
- Baethge, Christopher. 2012. „Nichtperfekter Impact-Faktor.“ *Deutsches Ärzteblatt* 109 (15): 267–69. <https://doi.org/10.3238/arztebl.2012.0267>.
- Besançon, Lonni, Guillaume Cabanac, Cyril Labbé, and Alexander Magazinov. 2024. „Sneaked References. Fabricated Reference Metadata Distort Citation Counts.“ *Asso for Info Science & Tech* 6 (2): 48–59. <https://doi.org/10.1002/asi.24896>.
- Bornmann, Lutz, Rüdiger Mutz, and Hans-Dieter Daniel. 2008. „Are There Better Indices for Evaluation Purposes Than the H Index? A Comparison of Nine Different Variants of the H Index Using Data from Biomedicine.“ *J Am Soc Inf Sci Tec* 59 (5): 830–37. <https://doi.org/10.1002/asi.20806>.
- Bornmann, Lutz, and Werner Marx. 2014. „Distributions Instead of Single Numbers: Percentiles and Beam Plots for the Assessment of Single Researchers.“ *J Assn Inf Sci Tec* 65 (1): 206–8. <https://doi.org/10.1002/asi.22996>.
- Cardoso, Fernando H. and Enzo Faletto. 1979. „Dependency and development in Latin America.“ 8th ed. Berkeley: University of California Press.
- Centrum voor Wetenschaps- en Technologiestudies (CWTS). 2024. „CWTS Leiden Ranking Open Edition. Data.“ Accessed January 26, 2025. <https://open.leidenranking.com/information/data>.
- Coalition for Reforming Research Assessment. 2022. „Agreement on Reforming Research Assessment.“ Accessed January 26, 2025. https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf.

- Cortegiani, Andrea, Filippo Sanfilippo, Jacopo Tramarin, and Antonino Giarratano. 2019. "Predatory Open-Access Publishing in Critical Care Medicine." *Journal of critical care* 50: 247–49. <https://doi.org/10.1016/j.jcrc.2018.12.016>.
- Culbert, Jack, Anne Hobert, Najko Jahn, Nick Haupka, Marion Schmidt, Paul Donner, and Philipp Mayr. 2024. "Reference Coverage Analysis of OpenAlex Compared to Web of Science and Scopus." Accessed November 19, 2024. <http://arxiv.org/pdf/2401.16359v3>.
- Deutsche Forschungsgemeinschaft. 2019. „Leitlinien Zur Sicherung Guter Wissenschaftlicher Praxis (Kodex).“ Bonn. Accessed January 26, 2025. <https://www.dfg.de/de/grundlagen-themen/grundlagen-und-prinzipien-der-foerderung/gwp/kodex>.
- Donner, Paul. 2024. „Data Inaccuracy Quantification and Uncertainty Propagation for Bibliometric Indicators.“ *Research Evaluation* 33:235. <https://doi.org/10.1093/reseval/rvae047>.
- DORA. 2012. „San Francisco Declaration on Research Assessment.“ Accessed January 26, 2025. <https://sfdora.org/read/>.
- Egghe, Leo. 2007. „From H to G: The Evolution of Citation Indices.“ *Research Trends* (5). http://info.scopus.com/resaerchtrends/archive/RT1/exp_op_1.htm.
- Garfield, Eugene. 1955. „Citation Indexes for Science. A New Dimension in Documentation Through Association of Ideas.“ *Science* 122 (3159): 105–11.
- Harzing, Anne-Wil, and Satu Alakangas. 2016. „Google Scholar, Scopus and the Web of Science: A Longitudinal and Cross-Disciplinary Comparison.“ *Scientometrics* 106 (2): 787–804. <https://doi.org/10.1007/s11192-015-1798-9>.
- Haustein, Stefanie. 2010. "Multidimensionale Zeitschriftenevaluation." In *Proceedingsband*, edited by Bernhard Mittermaier, 109-122. Konferenz der Zentralbibliothek, Forschungszentrum Jülich 5. Jülich: Forschungszentrum Jülich Zentralbibliothek.
- Haustein, Stefanie. 2023. „Stop Using the H-Index: Educating Researchers About Scholarly Metrics.“ March 28. Accessed August 26, 2024. https://uofc.sharepoint.com/sites/Knowledge-Engagement/_layouts/15/stream.aspx?id=%2Fsites%2FKnowledge%2DEngagement%2FNavigating%20DORA%2F2023%2D03%2D28%5FNavigatingDORA%5FStefanieHaustein%2FRecording%5FNavigatingDORA%5FHaustein%5F2023%2D03%2D28%2Emp4&ga=1.
- Hicks, Diana, Paul F. Wouters, Ludo Waltman, Sarah de Rijcke, and Ismael Rafols. 2015. „The Leiden Manifesto for Research Metrics.“ *Nature* 520: 429–31. <https://doi.org/10.1038/520429a>.
- Hirsch, Jorge Eduardo. 2005. „An Index to Quantify an Individual’s Scientific Research Output.“ *PNAS* 102 (46): 16569–72. <https://doi.org/10.1073/pnas.0507655102>.
- Koltun, Vladlen, and David Hafner. 2021. „The H-Index Is No Longer an Effective Correlate of Scientific Reputation.“ *PloS ONE* 16 (6): e0253397. <https://doi.org/10.1371/journal.pone.0253397>.
- Lesch, Harald. 2024. „Klima und Energie - Der Stand der Dinge.“ Stuttgart. Accessed January 26, 2025. https://www.youtube.com/watch?v=xWWXD_NKpHs.

- Liu, Weishu. 2021. „A Matter of Time: Publication Dates in Web of Science Core Collection.“ *Scientometrics* 126 (1): 849–57. <https://doi.org/10.1007/s11192-020-03697-x>.
- Nielsen, Michael. 2008. „Why the H-Index Is Little Use.“ Accessed August 04, 2016. <http://michaelnielsen.org/blog/why-the-h-index-is-virtually-no-use/>.
- OurResearch. 2024. “OpenAlex Technical Documentation, Sources.” Accessed January 26, 2025. <https://docs.openalex.org/api-entities/sources>.
- Rehs, Andreas. 2021. „A Supervised Machine Learning Approach to Author Disambiguation in the Web of Science.“ *Journal of Informetrics* 15 (3): 101166. <https://doi.org/10.1016/j.joi.2021.101166>.
- Richtig Georg, Marina Berger, Max Koeller, Markus Richtig, Erika Richtig, Jörg Scheffel, Marcus Maurer, and Frank Siebenhaar. 2023. “Predatory Journals: Perception, Impact and Use of Beall’s List by the Scientific Community—A Bibliometric Big Data Study.” *PloS ONE* 18 (7): e0287547. <https://doi.org/10.1371/journal.pone.0287547>.
- Reitz, Joan M. 2021. „Online Dictionary for Library and Information Science - Bibliometrics.“ Accessed July 12, 2021. https://www.abc-clio.com/ODLIS/odlis_b.aspx.
- Seglen, Per O. 1997. „Why the Impact Factor of Journals Should Not Be Used for Evaluating Research.“ *BMJ* 314:497. <https://doi.org/10.1136/bmj.314.7079.497>.
- Stringer Michael J., Marta Sales-Pardo, and Luís A. Nunes Amaral. 2010. “Statistical Validation of a Global Model for the Distribution of the Ultimate Number of Citations Accrued by Papers Published in a Scientific Journal.” *Journal of the American Society for Information Science and Technology* 61 (7): 1377–85. <https://doi.org/10.1002/asi.21335>.
- Waltman, Ludo, and Nees Jan P. van Eck. 2012. „The Inconsistency of the H-Index.“ *J Am Soc Inf Sci Tec* 63 (2): 406–15. <https://doi.org/10.1002/asi.21678>.
- Waltman, Ludo, Nees Jan P. van Eck, Martijn S. Visser, Mark Neijssel, Lucy Montgomery, Cameron Neylon, Bianca Kramer, Kyle Demes, and Jason Priem. 2024. “Introducing the Leiden Ranking Open Edition.” Accessed January 26, 2025. <https://www.leidenmadtrics.nl/articles/introducing-the-leiden-ranking-open-edition>.
- Yong, Alexander. 2014. „A Critique of Hirsch’s Citation Index. A Combinatorial Fermi Problem.“ *Notices Amer. Math. Soc.* 61 (9): 1040–50. <https://doi.org/10.1090/noti1164>.
- Zhu, Junwen, Guangyuan Hu, and Weishu Liu. 2019. „DOI Errors and Possible Solutions for Web of Science.“ *Scientometrics* 118 (2): 709–18. <https://doi.org/10.1007/s11192-018-2980-7>.