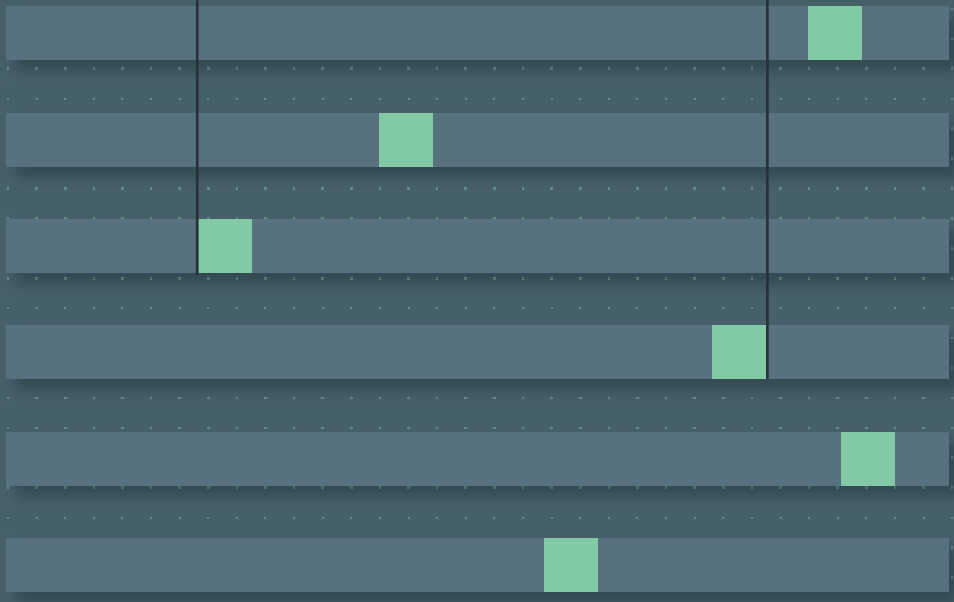
STUTTGART UNIVERSITY PROCEEDINGS

VERMESSENE PHILOSOPHIE

KONSTRUKTION UND KONTROLLE WISSENSCHAFTLICHER EXPERTISE IM DIGITALEN RAUM



NICOLA MÖSSNER & KLAUS ERLACH (HRSG.)

Vermessene Philosophie

Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum

Nicola Mößner und Klaus Erlach (Hrsg.)

]u[

ubiquity press
London

Published by
Ubiquity Press Ltd.
Unit 3N, 6 Osborn Street
London, E1 6TD
United Kingdom
www.ubiquitypress.com

Text © Nicola Mößner and Klaus Erlach 2026

First published 2026

Supporting Agency
Diese Publikation wurde durch das Bundesministerium
für Forschung, Technologie und Raumfahrt (BMFTR)
im Rahmen des Projekts Diamond_Thinking gefördert.

Cover design by Nadine Müller
Cover image credit by Nadine Müller

Print and digital versions typeset by Siliconchips Services Ltd.

ISBN (Paperback): 978-1-914481-78-9
ISBN (PDF): 978-1-914481-79-6
ISBN (EPUB): 978-1-914481-80-2
ISBN (Mobi): 978-1-914481-81-9

DOI: https://doi.org/10.5334/bdh

This work is licensed under the Creative Commons Attribution 4.0 International License (unless stated otherwise within the content of the work). To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/ or send a letter to Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA. This license allows for copying any part of the work for personal and commercial use, providing author attribution is clearly stated.

The full text of this book has been peer-reviewed to ensure high academic standards. For full review policies, see http://www.ubiquitypress.com/

Suggested citation:
Mößner, N. und Erlach, K. (Hrsg.) 2026. *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.
License: CC BY 4.0

To read the free, open access version of this book online, visit https://doi.org/10.5334/bdh or scan this QR code with your mobile device:

Inhalt

1. Einleitung: *verTRACKte Infrastruktur?!*
Nicola Mößner & Klaus Erlach 1

Teil I: Die Praxis des elektronischen Publizierens und ihre Probleme

Teil I: Die Praxis des elektronischen Publizierens und ihre Probleme
Klaus Erlach & Nicola Mößner 17

2. Verbreitung wissenschaftlicher Erkenntnisse durch profitorientierte oder öffentliche digitale Infrastruktur?
Eric Retzlaff 23

3. Diamond Open Access?
Stefan Drößler 43

4. Datentracking – Gefahr für die Wissenschaft?
Bernhard Mittermaier 69

5. Transparenz statt Transzendenz: Orientierung im ABC der Publikationskosten
Lisa-Marie Stein 97

Teil II: Die Frage nach der Expertise

Teil II: Die Frage nach der Expertise
Klaus Erlach & Nicola Mößner 131

6. Peer Review, Bibliometrie, Altmetrics – gibt es einen Königsweg bei Evaluierungen?
Bernhard Mittermaier 139

7. Wenn die Auswahl der Experten Expertise erfordert:
Die Debatte um den Ursprung von SARS-CoV-2
und das pervasive Verhältnis von Vertrauen und Wissen
Andreas Kaminski & Jörn Wiengarn 169

8. Vom Zählen zum Bewerten – das Wunder objektiver
Wissenschaftler-Beobachtung
Klaus Erlach 223

9. Epistemic Criteria for Peer Review in Philosophy of Science:
A Qualitative Inquiry into a Social Epistemic Practice
Nora Hangel 247

Teil III: Auswirkungen auf die philosophische Praxis und Lösungsvorschläge

Teil III: Auswirkungen auf die philosophische Praxis
und Lösungsvorschläge
Nicola Mößner & Klaus Erlach 287

10. Perils and Promises: How to Regain Control
of Metrics in Academia
Hanne Andersen 295

11. Open Science – eine Frage der Wissenschaftskultur
Nicola Mößner 319

12. Positionspapier: „Elektronisches Publizieren
und Bewertung wissenschaftlicher Expertise“
Nicola Mößner & Klaus Erlach 353

13. Das Projekt Diamond_Thinking
Stefan Drößler 367

Verzeichnis der AutorInnen 375

KAPITEL I

Einleitung: *verTRACKte Infrastruktur?!*

Nicola Mößner und Klaus Erlach

Es kriselt im Wissenschaftsbetrieb, wie man derzeit allerorten hört. Neben die Publikations-, die Subskriptions- und die Replikationskrise (vgl. z.B. Brembs u.a. 2023) tritt seit Neuestem noch eine weitere Krise hinzu, nämlich die sogenannte Überwachungs- oder Tracking-Krise (vgl. z.B. Siems 2022; 2023; Mößner 2022b; 2023b). Sie alle sind aufs Engste verbunden mit den Prozessen des elektronischen Publizierens, welche zusammen mit ihren Auswirkungen im folgenden Band kritisch in den Blick genommen werden sollen.

Wissenschaftliche Ergebnisse – auch jene der akademischen Philosophie – werden heutzutage meist digital veröffentlicht. Die gedruckte Fachzeitschrift hat so gut wie ausgedient. Auch das gedruckte Buch tut sich zunehmend schwer, im Diskurs der Fachwissenschaften seinen Platz als Medium der Speicherung und Verteilung von Wissen zu bewahren. Der übliche Weg der wissenschaftlich relevanten Veröffentlichungen führt derweil über elektronische Zeitschriften.

Zitationsvorschlag für dieses Kapitel:
Mößner, N. und Erlach, K. 2026. Einleitung: *verTRACKte Infrastruktur?!*. In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* S. 1–13. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.a.
License: CC BY 4.0

Verlage präsentieren ihre Angebote mittels Webinterface online. Datenbanken halten die Artikel für Wissenschaftler:innen und deren Recherchen zugänglich und verfügbar.

Die Frage des Zugangs bzw. der Zugänglichkeit ist dabei einer der Aspekte, der besagte Krisen befeuert. Sogenannte Subskriptionsmodelle stehen im Fokus. Der einzelne Forschende merkt zumeist wenig von den damit verbundenen Schwierigkeiten, denn es sind die Bibliotheken, die für die Angehörigen ihrer Institutionen den Zugang zu den Informationsangeboten sicherstellen müssen und somit als Verhandlungspartner für die großen Wissenschaftsverlage agieren. Als Mittler zwischen Angebot (Verlagshäuser) und Nachfrage (Wissenschaftler:innen) sind Vertreter:innen aus dem Bibliothekswesen in einer marktwirtschaftlichen Zwickmühle gefangen: nur schwerlich können sie die Preispolitik der Anbieter wirksam in Frage stellen, wenn sie ihrem Auftrag der Informationsversorgung bestmöglich nachkommen wollen. Hier setzt die sogenannte *Subskriptionskrise* an, zu der die Deutsche Forschungsgemeinschaft (DFG) in einem entsprechenden Positionspapier Folgendes festhält: „In den letzten Jahren zeichnete sich ein starkes Wachstum bei Subskriptionskosten und Gewinnmargen bei führenden Verlagsgruppen ab [...]. Preispolitik und Geschäftsmodelle wie z. B. die Paketbündelung von Zeitschriften und auch E-Books erschwerten es der Wissenschaft zunehmend, bezahlbaren Zugriff auf alle relevanten wissenschaftlichen Inhalte sicherzustellen" (Deutsche Forschungsgemeinschaft, AG Publikationswesen 2022: 29).

Kosten fallen mittlerweile längst nicht mehr allein für das Abonnement von Lesezugriffen auf Zeitschriftenartikel an. Das Zeitalter des elektronischen Publizierens hat es ermöglicht, das Veröffentlichen dieser Artikel selbst kostenpflichtig werden zu lassen. So werden für die Publikation von Zeitschriftenbeiträgen im sogenannten Gold-Open-Access-Modell weitere Gebühren erhoben, die unter der Bezeichnung ‚*Article Processing Charges* (APCs)' bekannt geworden sind. Auf dem Zeitschriftenmarkt zeichnet sich in diesem Zusammenhang eine starke Oligopolbildung der Anbieter ab (Drößler 2022: 91ff.).

Verstärkt wird diese Zentralisierung ferner dadurch, dass die Oligopolisten bestrebt sind, den gesamten Forschungszyklus in den Wissenschaften mit ihren Produkten abzudecken. Von der Recherche über die Publikation bis zur Evaluation könne künftig die gesamte Forschungsarbeit über entsprechende elektronische Plattformen (z.B. SCOPUS von Elsevier) abgewickelt werden, so das Versprechen der Unternehmen.

Die zugehörigen Anbieter stellen sich zwischenzeitlich als Informationsanalysten vor, nicht mehr als bloße Verlagshäuser (vgl. https://www.elsevier.com/about, eingesehen am 21. Juli 2025). Sie machen damit öffentlich, dass ihr Geschäftsmodell weit über das klassische Verständnis der Arbeit eines Wissenschaftsverlags hinausreicht. Ihr Geschäft sind vielmehr die Daten und Datenspuren der Wissenschaftler:innen auf ihren Plattformen (vgl. Lamdan 2023). Ebenso wie die Anbieter von Suchmaschinen oder Webshops im Internet betreiben sie das zumeist verdeckte Sammeln von sowohl öffentlich zugänglichen als auch nicht-öffentlich zugänglichen Daten ihrer Nutzer:innen und die Erstellung entsprechender Profile von diesen, die dann weiter vermarktet werden können. Dies ist der Grundmechanismus der Überwachungs- bzw. *Trackingkrise*; eine Krise, deren Tragweite betroffene Wissenschaftler:innen und ihre Institutionen gerade erst zu erfassen beginnen (vgl. DFG-Ausschuss für wissenschaftliche Bibliotheken und Informationssysteme 2021).

Noch seltener wird gesehen, dass diese Entwicklung letztlich nur möglich ist, weil die Wissenschaft sich noch in einer weiteren Hinsicht (neben der bequemen digitalen Arbeitspraxis, die alle Leistungen aus einer Hand – d.h. von einem Zugang aus verspricht) vom sprichwörtlichen Tropf der Data-Analytics-Unternehmen abhängig gemacht hat: Erst die angesprochenen Datenbanken im Hintergrund des elektronischen Publizierens erlauben es nämlich, jene Kennzahlen zu erheben, welche für die bibliometrische Analyse und Bewertung wissenschaftlicher Expertise heutzutage eine so große Rolle spielen. Die Auswertung von Zitationen gilt nach wie vor als Goldstandard für die Erfassung des (vermeintlichen) Impacts der Forschenden. Auch

der sogenannte Journal Impact-Faktor basiert auf diesen Zahlenwerten und wird als Maßstab für die Reputation eines Publikationsortes gewertet. Metriken der genannten Art setzen also die Erfassung von Veröffentlichungen in Datenbanken voraus, welche sich wiederum in den Händen besagter Oligopolisten (derzeit insbesondere *Wiley, Springer / Nature, Elsevier*) befinden. Das Reputationsdenken der Wissenschaft, das sich nach wie vor gern mit vermeintlich objektiven Zahlenwerten schmückt, ist also ein treibender Faktor dafür, dass die beteiligten Unternehmen einen stetig wachsenden Einfluss im Wissenschaftssektor für sich verzeichnen können.

Hieran knüpft dann eine weitere Krise an, nämlich die sogenannte *Publikationskrise* (vgl. Mößner 2023a; Deutsche Forschungsgemeinschaft, AG Publikationswesen 2022; Andersen 2019). Die Forschenden sind sich natürlich bewusst, welcher Wert den angesprochenen Kennzahlen in ihrer Community zugesprochen wird – sei es im Kontext von Bewerbungsverfahren, sei es bei der Bewilligung von Forschungsstipendien oder Fördergeldern etc. Wissenschaftliches Publizieren wird daher immer häufiger von karriere-taktischen Erwägungen bestimmt, die wiederum zu Lasten der Normen guter wissenschaftlicher Arbeitsweise gehen können. Dazu zählen beispielsweise Überlegungen, Forschungsergebnisse in möglichst kleinen ‚Häppchen' zu veröffentlichen, um eine maximale Anzahl an zitierfähigen Publikationen zu erzielen und anderes mehr. Wissenschaftliche Veröffentlichung sollten aber dem Zweck dienen, Inhalte zu verbreiten und anderen Forschenden zugänglich zu machen, um das kollektive Unternehmen der Wissenschaft voranzubringen. Eine andere Schwierigkeit stellen Zitationskartelle dar, in denen Zitate auf der Basis persönlicher Bekanntschaften unter den Autor:innen oder aus Gefälligkeiten heraus vergeben werden. Auf diese Weise können Zitationszahlen und damit die vermeintliche Reputation eines Forschenden gesteigert werden, ohne dass dies mit einer tatsächlichen Leistungsverbesserung einherginge.

Schließlich wird seit einiger Zeit auch über die sogenannte *Replikationskrise* diskutiert (vgl. Brembs u.a. 2023; Fidler und Wilcox

2018). Eine basale Annahme in den Wissenschaften besagt, dass Forschungsergebnissen und -hypothesen prinzipiell intersubjektiv überprüfbar sein sollen. Was aber bedeutet es, wenn sich herausstellt, dass viele Experimente und damit die durch sie gewonnenen Ergebnisse gar nicht wiederholt werden (können)? Wiederum kann das System des elektronischen Publizierens eine Rolle bei derlei Fehlschlägen spielen. Verleitet es durch seine Verknüpfung mit der Metrisierung wissenschaftlicher Expertise doch eventuell zur vorschnellen Veröffentlichung von (vermeintlichen) Ergebnissen, die dann eben auch nicht reproduzierbar sind.

All diese krisenhaften Phänomene, welche auf die eine oder andere Weise an das elektronische Publizieren und die damit einhergehenden Entwicklungen gebunden sind, können sich nun in mehrfacher Hinsicht nachteilig im Kontext der Wissenschaft und auch im Hinblick auf die Relation zwischen Wissenschaft und Gesellschaft auswirken. Die angeführten Krisen betreffen gleichermaßen die Generierung als auch die Verbreitung wissenschaftlicher Erkenntnis. In beiden Bereichen kann es in der Folge zu dysfunktionalen Entwicklungen kommen, wie am Beispiel der Publikationskrise schon angedeutet wurde. Steht nicht mehr der gemeinsame Erkenntnisgewinn im Fokus, sondern das egozentrische Streben nach Impact-Punkten, dann dominieren taktische Publikationserwägungen statt der Wahrheitssuche usw.

Wird der kennzahlengetriebene interne Mechanismus des Wissenschaftsbetriebs aber für die Öffentlichkeit – insbesondere für wissenschaftliche Laien – sichtbar, kann dies das Vertrauen in wissenschaftliche Expertise zusätzlich unterminieren (vgl. Mößner 2022a; Retzlaff 2022; Oreskes 2019). Für den Außenstehenden stellt sich das wissenschaftliche Arbeiten dann eher als eine Profilierungstätigkeit dar, denn als die Arbeit an der Lösung drängender gesellschaftlicher Probleme. Hat man aber erst einmal den Eindruck, dass das Streben nach Erkenntnis hinter dem Sammeln von Zitaten und anderen Reputationsfaktoren zurücktritt, liegt eine misstrauische Haltung gegenüber wissenschaftlichen Ergebnissen, deren Geltungsanspruch sowie den auf ihnen basierenden Empfehlungen und Handlungsentscheidungen in anderen gesellschaftlichen

Bereichen (z.B. Politik, Bildungswesen, Wirtschaft) nahe. Ein solcher Zweifel an wissenschaftlicher Expertise wurde z.B. besonders deutlich während der Jahre der Corona-Pandemie (vgl. die Beiträge in Rico Hauswald und Pedro Schmechtig 2023).

Schließlich steht mit dem wachsenden Einfluss der Data-Analytics-Unternehmen die Autonomie der Wissenschaft selbst zur Disposition. Dies kann sowohl auf der Ebene des individuellen Forschenden eine Rolle spielen, als auch für einzelne Institutionen sowie für die Wissenschaft insgesamt. Für den einzelnen kann dies bedeuten, dass seine Forschungsinteressen zu einem Nischenthema werden, für das keine Fördergelder etc. mehr zu erhalten sind. Für Institutionen kann es heißen, dass sie ihre Forschungsprofile insgesamt auf Themen ausrichten, die auf Trendanalysen der Datenbankbetreiber beruhen. Die Vorhersagen konkreter Forschungsthemen können einerseits nur so zuverlässig sein, wie es die Validität der genutzten Datenbasis zulässt. Andererseits droht hier, in einem selbstverstärkenden Prozess einseitige Mainstream-Forschung zu entstehen.

Schließlich geht mit einer derart auf „predictive" und „prescriptive data" gestützten Wissenschaft (vgl. Holzer 2022; Herb 2018) auch die Gefahr des Verlusts gesellschaftspolitischer Relevanz und Legitimation der Forschungstätigkeit einher. Wenn Forschungsfragen nicht mehr ihren Ursprung im gemeinsamen Wissensstreben haben, sondern als Vorgaben aus einer (nach wie vor) unvollständigen Datenbasis auf den Plattformen besagter Informationsanalysten automatisch generiert werden, kann vermutet werden, dass die Verbindung zwischen Öffentlichkeit und Wissenschaftler:innen weiter geschwächt wird. Schon heute wird von verschiedenen Seiten auf ein zunehmend prekäres Verhältnis der beiden Seiten aufmerksam gemacht. Es wird angemerkt, dass das Vertrauen in wissenschaftliche Erkenntnis auch deshalb schwinde, weil wissenschaftliche Laien (zumindest in ihrer Wahrnehmung) nur geringe Mitsprachemöglichkeiten bei der Wahl von Forschungsthemen haben. Sie haben den Eindruck, dass Wissenschaft nicht im Sinne des öffentlichen Wohls betrieben wird, sondern im Interesse elitärer Zirkel im Elfenbeinturm weit jenseits

gesellschaftlicher Bedarfe und Probleme. Als Folge davon wenden sich bestimmte Gruppen von klassischen Expert:innen und deren Ratschlägen und Empfehlungen ab (vgl. Mößner 2024; Oreskes 2019; Nichols 2017; Oreskes und Conway 2012). Die beschriebenen Mechanismen des elektronischen Publizierens weisen nun das Potential auf, diese Schwierigkeiten weiter zu verstärken.

Obwohl diese Entwicklungen bereits von verschiedener Seite erkannt und Änderungen angemahnt wurden (vgl. Altschaffel u.a. 2024 ; Brembs u.a. 2023; 2020), scheint der Wissenschaftsbetrieb in den eingefahrenen Bahnen zu verharren und diese auch weiter auszubauen. Im vorliegenden Band soll diesem Phänomen genauer auf den Grund gegangen werden. Insbesondere soll dem Zusammenhang zwischen den aufgezeigten dysfunktionalen Entwicklungen im Bereich des elektronischen Publizierens und dem auf Metriken beruhenden Reputationsdenken in den Wissenschaften nachgespürt werden. Es wird deutlich, dass hier in der Tat eine enge Verbindung besteht, die zumindest mitverantwortlich dafür ist, dass das bestehende System des elektronischen Publizierens weiterhin unkritisch adapiert wird, inklusive der Herausbildung von Oligopolisten auf dem Zeitschriftenmarkt und der Weiterentwicklung von Wissenschaftsverlagen zu Data-Analytics-Unternehmen, die mit dem verdeckten Datensammeln und der Profilerstellung ihrer Nutzer:innen weiteren monetären Vorteil aus dem Wissenschaftssystem ziehen.

Neben der Analyse der Problemstellung wird mit den Beiträgen des vorliegenden Bandes darüber hinaus das Ziel verfolgt, zumindest für einige Bereiche Vorschläge für mögliche Veränderungen und Verbesserungen vorzulegen. Leitende Fragestellungen sind dabei die folgenden:

a) Wie kann die wissenschaftliche Gemeinschaft ihre Autonomie gegenüber einem Oligopol der Informationsanalysten behaupten?
b) Wie können beunruhigende Fehlentwicklungen wie jene des Wissenschaftler-Trackings rückgängig gemacht werden?
c) Welche Kriterien und Maßstäbe sollte die wissenschaftliche Gemeinschaft an ihre eigene Expertisebeurteilung künftig stellen?

Für die Auswahl und Zusammenstellung der vorliegenden Beiträge war der Gedanke zentral, dass der Philosophie hier eine besondere Bedeutung zukommt, weil sie in doppelter Hinsicht mit den aufgeworfenen Fragestellungen konfrontiert ist (Abb. 1.1). So sind Philosoph:innen einerseits diejenigen, die auf einer theoretischen Ebene über die genannten Phänomene reflektieren. Hier geht es um die Fragen, was überhaupt einen Experten ausmacht, was Expertise kennzeichnet, wie man einen Experten / eine Expertin erkennen kann sowie ebenso darum zu klären, was Vertrauen in wissenschaftliche Expertise ausmacht, was sie schwächen bzw. stärken kann, welche Aufgaben der Wissenschaft innerhalb der Gesellschaft zukommen (sollen) und wie die Schnittstelle zwischen Wissenschaft und den anderen gesellschaftlichen Subsystemen künftig beschaffen sein sollte.

Andererseits sind Philosoph:innen als Angehörige der akademischen Institution selbst von den beschriebenen Entwicklungen und Problemen betroffen. So stellt die Auseinandersetzung mit der Frage, was wir unter ‚wissenschaftlicher Expertise' aktuell – unserer wissenschaftlichen Praxis zufolge – verstehen und was wir darunter eigentlich – gemäß unserer wissenschaftlichen Ideale – verstehen wollen, letztlich eine selbstreflexive Aufgabe für die Philosophie als akademische Disziplin dar. Wie möchten wir als Philosoph:innen unsere wissenschaftliche Leistung bewertet sehen und nach welchen Maßstäben? Ebenso müssen Philosoph:innen sich mit den Mechanismen und Folgen des elektronischen Publizierens in ihrem eigenen Fach auseinandersetzen. Welche alternativen Wege könnten wir hier einschlagen, um den aufgezeigten Schwierigkeiten der Abhängigkeit von Oligopolisten und Data-Analytics-Unternehmen zu entgehen? Zusammenhängend damit stellt sich auch für die akademische Philosophie die Frage nach ihrer gesellschaftlichen Glaubwürdigkeit.

Die Philosophie tritt somit sowohl als Reflexions- als auch als Anwendungswissenschaft im Umfeld der erörterten Fragestellungen auf (vgl. Abb. 1.1).

Der folgende Band versammelt aber nicht allein Beiträge aus der Philosophie, sondern ebenso solche aus den Bibliotheks-



Abbildung 1.1: Die Rolle der Philosophie.

und Informationswissenschaften, deren Vertreter:innen sich viel unmittelbarer mit den Mechanismen und Problemen des elektronischen Publizierens konfrontiert sehen, als viele Wissenschaftler:innen dies in ihrem Tagesgeschäft erleben. Diese Perspektive mit einzubringen, kann damit in einem Bereich zu mehr Transparenz beitragen, der zumindest im kommerziellen Umfeld oftmals stark von einer bewusst aufrechterhaltenen Intransparenz profitiert.

Die Autor:innen des Bandes diskutieren in ihren Beiträgen die Themenstellungen in den Spannungsfeldern von Autonomie und Expertise einerseits sowie von wissenschaftlicher Praxis und Evaluationsdruck andererseits. Was könnten zeitgemäße Distributionswege wissenschaftlicher Erkenntnis sein?

Der Band gliedert die Beiträge der Autor:innen in drei Teile, die drei Reflexionsebenen entsprechen (vgl. Abb. 1.2). Er beginnt auf der (mittleren) Objektebene des Wissenschaftssystems mit der unmittelbaren Betrachtung der mit dem elektronischen

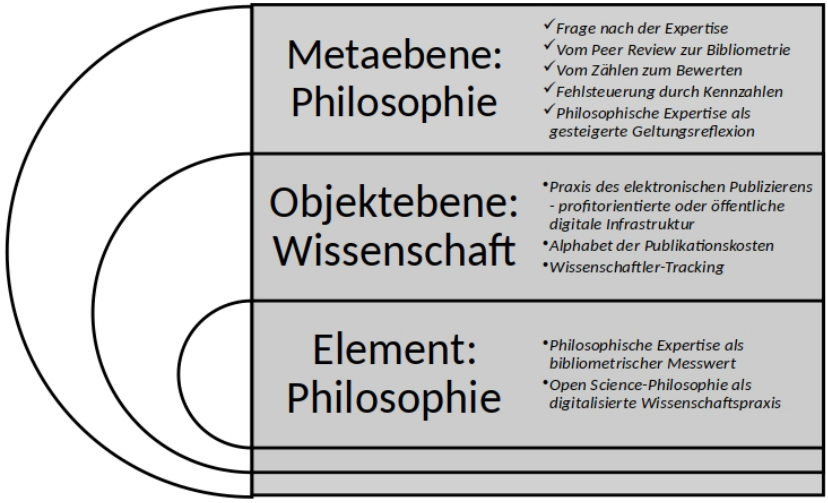


Abbildung 1.2: Reflexionsebenen.

Publizieren verbundenen Herausforderungen, Problemen und Auswirkungen. (Teil I: *Die Praxis des elektronischen Publizierens und ihre Probleme*). Anschließend folgt auf der (oberen) Metaebene der Philosophie als Reflexionswissenschaft die Betrachtung von Expertise, Vertrauen und Evaluierung (Teil II: *Die Frage nach der Expertise*). Und abschließend folgt auf der (unteren) elementaren Ebene der Philosophie als Anwendungswissenschaft die Entwicklung der kompakt in einem Positionspapier festgehaltenen Lösungsvorschläge für die Philosophie als Wissenschaft (Teil III: *Auswirkungen auf die philosophische Praxis und Lösungsvorschläge*).

Eine Reihe der Beitragenden hat die gemeinsame Diskussion der Themen bereits beim Workshop „verTRACKte Infrastruktur?! Von Konstrukteuren und Kontrolleuren wissenschaftlicher Expertise“ im September 2023 an der Württembergischen Landesbibliothek in Stuttgart begonnen. Die Diskussionsergebnisse des Workshops sind in die Beiträge des vorliegenden Bandes eingeflossen. Darüber hinaus haben die Herausgebenden im Zuge ihres Reviewprozesses den argumentativen Aufbau des Bandes gestaltet, ergänzende Beiträge eingeworben sowie Zusammenhänge zwischen den einzelnen Teilen transparent gemacht.

Die Herausgebenden dieses Bandes möchten an dieser Stelle noch einmal ihren Dank an die Förderer der genannten Veranstaltung aussprechen, nämlich der Deutschen Gesellschaft für Philosophie e.V. (DGPhil) und dem Internationalen Zentrum für Kultur- und Technikforschung (IZKT) der Universität Stuttgart. Bedanken möchten wir uns auch bei allen Teilnehmenden des Workshops sowie bei den übrigen Beitragenden des vorliegenden Bandes für den produktiven Austausch und die Ideen für eine bessere akademische Zukunft.

Diese Publikation wurde durch das Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) im Rahmen des Projekts Diamond_Thinking (Förderkennzeichen:16KOA006A) gefördert. Auch hierfür möchten wir unseren Dank ausdrücken.

Literatur

Altschaffel, Robert, Michael Beurskens, Jana Dittmann, u.a. (2024) „Datentracking und DEAL – Zu den Verhandlungen 2022/2023 und den Folgen für die wissenschaftlichen Bibliotheken". *RuZ – Recht und Zugang* 5 (1): 23–40. https://doi.org/10.5771/2699-1284-2024-1-23.

Andersen, Hanne (2019) „Can Scientific Knowledge Be Measured by Numbers?" In *What Is Scientific Knowledge?: An Introduction to Contemporary Epistemology of Science*, hrsg. von Kevin McCain und Kostas Kampourakis. Routledge, S. 144–159. https://doi.org/10.4324/9780203703809.

Brembs, Björn, Konrad Förstner, Peter Kraker, u.a. (2020) „Auf einmal Laborratte". *Zenodo*. https://doi.org/10.5281/ZENODO.4317253.

Brembs, Björn, Philippe Huneman, Felix Schönbrodt, u.a. (2023) „Replacing Academic Journals". *Zenodo.* https://doi.org/10.5281/ZENODO.7974116.

Deutsche Forschungsgemeinschaft, AG Publikationswesen (2022) „Wissenschaftliches Publizieren als Grundlage und Gestaltungsfeld der Wissenschaftsbewertung". *Zenodo*. https://doi.org/10.5281/ZENODO.6538163.

DFG-Ausschuss für wissenschaftliche Bibliotheken und Informationssysteme (2021) „Datentracking in der Wissenschaft: Aggregation

und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage. Ein Informationspapier des Ausschusses für Wissenschaftliche Bibliotheken und Informationssysteme der Deutschen Forschungsgemeinschaft". *Zenodo*. https://doi.org/10.5281/ZENODO.5900759.

Drößler, Stefan (2022) „Open Access: Technikgetriebene Gesellschaftsutopie für die Transformation des wissenschaftlichen Publikationssystems". In *Kalibrierung der Wissenschaft: Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis*, hrsg. von Nicola Mößner und Klaus Erlach. Transcript Verlag, S. 79–101. https://doi.org/10.14361/9783839462102-004.

Fidler, Fiona und John Wilcox (2018) „Reproducibility of Scientific Results". In *The Stanford Encyclopedia of Philosophy*, Winter 2018, hrsg. von Edward N. Zalta. https://plato.stanford.edu/archives/win2018/entries/scientific-reproducibility/.

Hauswald, Rico und Pedro Schmechtig (Hrsg.) (2023) *Wissensproduktion und Wissenstransfer in Zeiten der Pandemie. Der Einfluss der Corona-Krise auf die Erzeugung und Vermittlung von Wissen in wissenschaftstheoretischer, epistemologischer und bildungsphilosophischer Perspektive*. Karl Alber.

Herb, Ulrich (2018) „Zwangsehen und Bastarde. Wohin steuert Big Data die Wissenschaft?" *Information – Wissenschaft & Praxis* 69 (2–3): 81–88. https://doi.org/10.1515/iwp-2018-0021.

Holzer, Angela (2022) „Die Vermessbarkeit der Wissenschaft: Digitalisierung, wissenschaftliches Publizieren, Verhaltenstracking und Wissenschaftsbewertung". In *Kalibrierung der Wissenschaft: Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis*, hrsg. von Nicola Mößner und Klaus Erlach. Transcript Verlag, S. 163–182. https://doi.org/10.14361/9783839462102-007.

Lamdan, Sarah (2023) *Data cartels: the companies that control and monopolize our information*. Stanford University Press.

Mößner, Nicola (2022a) „Wissenschaft in ›Unordnung‹?: Gefiltertes Wissen und die Glaubwürdigkeit der Wissenschaft". In *Kalibrierung der Wissenschaft: Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis*, hrsg. von Nicola Mößner und Klaus Erlach. Transcript Verlag, S. 103–136. https://doi.org/10.14361/9783839462102-005.

Mößner, Nicola (2022b) „›Keine Spyware in Bibliotheken und Institutionen‹ – wirklich?: Die Gefahren des Wissenschaftstrackings". *Mitteilungen der DGPhil* 55 (Frühjahr): 11–13.

Mößner, Nicola (2023a) „Digitale Praxis – Fallstrick für Normen im Wissenschaftsalltag?“ *Prae|Faktisch. Ein Philosophieblog*. https://praefaktisch.de/digitalisierung/digitale-praxis-fallstrick-fuer-normen-im-wissenschaftsalltag/.

Mößner, Nicola (2023b) „Noch nicht ist schon zu viel“. *Jahrbuch Technikphilosophie. Faktor Mensch*, Vol. 9: 253–257. https://doi.org/10.5771/9783748941767-253.

Mößner, Nicola (2024) „Sucht oder Autonomie? – Neue ExpertInnen im Netz“. In *Philosophische Digitalisierungsforschung. Verantwortung, Verständigung, Vernunft, Macht*, hrsg. von Rainer Adolphi, Suzana Alpsancar, Susanne Hahn, und Matthias Kettner. Transcript, S. 197–217. https://doi.org/10.1515/9783839474976.

Nichols, Tom (2017) *The death of expertise the campaign against established knowledge and why it matters*. Oxford University Press.

*Oreskes, Naomi (*2019) *Why Trust Science?* Princeton University Press.

Oreskes, Naomi und Erik M. Conway (2012) *Merchants of Doubt. How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. Bloomsbury.

Retzlaff, Eric (2022) „Wer bewertet mit welchen Interessen wissenschaftliche Publikationen?: Eine Skizzierung des Einflusses kommerzieller Interessen auf die Forschungsoutput-Bewertung“. In *Kalibrierung der Wissenschaft: Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis*, hrsg. von Nicola Mößner und Klaus Erlach. Transcript Verlag, S. 139–162. https://doi.org/10.14361/9783839462102-006.

Siems, Renke (2022) „Das Lesen der Anderen. Die Auswirkungen von User Tracking auf Bibliotheken“. *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB*, 9(1): 1–25. https://doi.org/10.5282/O-BIB/5797.

Siems, Renke (2023) „‚Überwachen und Strafen‘ – Tracking und Kontrolle des Forschungszyklus“. *ABI Technik* 43 (2): 86–95. https://doi.org/10.1515/abitech-2023-0016.

TEIL I

Die Praxis des elektronischen Publizierens und ihre Probleme

TEIL I

Die Praxis des elektronischen Publizierens und ihre Probleme

Klaus Erlach und Nicola Mößner

Im Jahr 2021 wies die Deutsche Forschungsgemeinschaft (DFG) darauf hin, dass durch die Bündelung von zentralen Dienstleistungen und relevanter digitaler Infrastruktur im Bereich der Wissenschaftskommunikation in den Händen weniger großer internationaler Verlagskonzerne die Gefahr eines „Wissenschaftlertrackings" entstünde (vgl. DFG 2021). Kommerzielle Betreiber entsprechender IT-Lösungen (Datenbanken etc.) können umfangreiche Profile ihrer Nutzer:innen erstellen, die monetär nutzbar sind. Auf diese Weise wird der Verwertungskreislauf wissenschaftlicher Ergebnisse (siehe Abbildung) um eine fünfte Dimension zu Gunsten internationaler Konzerne erweitert: Sie profitieren insofern, als

(a) Wissenschaftler:innen ihnen alle Inhalte fertig aufbereitet gratis zur Verfügung stellen, um sie später als Leser:innen zu bezahlen,

Zitationsvorschlag für dieses Kapitel:
Erlach, K. und Mößner, N. 2026. Teil I: Die Praxis des elektronischen Publizierens und ihre Probleme. In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* S. 17–22. London: Ubiquity Press.
DOI: https://doi.org/10.5334/bdh.b. License: CC BY 4.0

(b) sie deren Begutachtung an die Wissenschaftler:innen zurückdelegieren und dadurch eine kostenfreie Qualitätssicherung erhalten,
(c) sie das fertige Publikationsprodukt den wissenschaftlichen Bibliotheken verkaufen und
(d) sie die dabei erzeugten bibliometrischen Daten über Autor:innen, Institutionen und Länder zusätzlich im Zuge der Erstellung von Rankings etc. vermarkten (vgl. Holzer 2022).

In der neuesten fünften Dimension verwerten die Verlage nun auch noch:

(e) die Datenspuren, die recherchierende und lesende Wissenschaftler:innen bei der Nutzung ihrer IT-Lösungen hinterlassen.

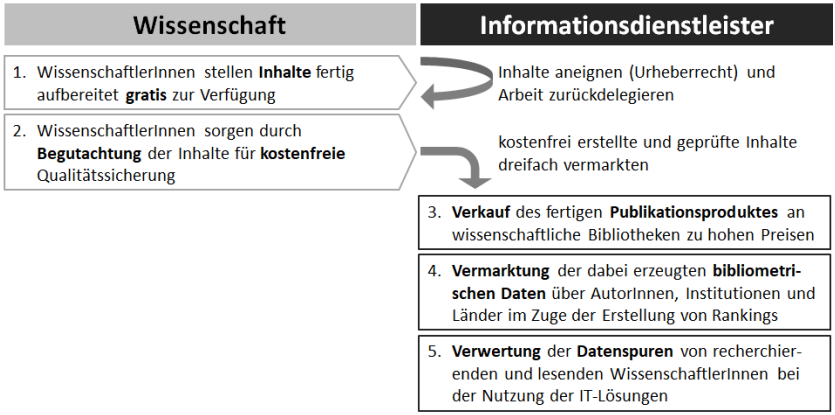


Abbildung: Das Geschäftsmodell der Verlagskonzerne im digitalen Wandel besteht darin, unentgeltlich von Wissenschaftler:innen bereitgestellte Ergebnisse kommerziell mehrfach zu verwerten (in Erweiterung zu Holzer 2022).

Zur Subskriptions- und Replikationskrise tritt damit zusätzlich eine Krise des Datenschutzes und verletzter Privatsphären.

Von Expertenseite wird bereits seit einiger Zeit auf diese Fehlentwicklung hingewiesen (vgl. z.B. Retzlaff 2022), welche derzeit in einer klaren Oligopolstellung bestimmter Informationsanalysten –

so die Selbstbeschreibung großer Verlagskonzerne in der heutigen Zeit (vgl. Elsevier 2022) – gipfelt. Auf Marktmechanismen zu setzen, um zu einer Ausbalancierung von Interessen und Angeboten zurückzufinden, ist somit nicht mehr möglich. Stattdessen gibt es einen klaren Trend zu steigenden Preisen und Gewinnmargen.

Die Bindung der wissenschaftlichen Community an die Konzerne hat darüber hinaus eine neue Intensität erlangt. Grund hierfür ist, dass sich ausschließlich durch die Nutzung der von diesen bereitgestellten Datenbanklösungen die im Wissenschaftsalltag zur Bewertung und Kennzeichnung vermeintlicher Expertise übergreifend herangezogenen bibliometrischen Daten und Kennzahlen (h-Index, Impact Faktor etc.) erheben lassen. Unreflektiert bleibt dabei der tatsächliche Aussagewert dieser Zahlen, sind sie doch stets nur auf eine ausgewählte Datenbasis bezogen, deren Repräsentanz, Vollständigkeit und Qualität weder sichergestellt noch für die wissenschaftliche Community transparent ist. *Datenanalysten werden somit zu Konstrukteuren und Kontrolleuren wissenschaftlicher Expertise.* Erhebliche Nebenwirkungen auf Form und Inhalte des wissenschaftlichen Arbeitens sind nicht ausgeschlossen und zugleich im Einzelnen nicht nachweisbar, allenfalls als statistischer Effekt.

Möglichkeiten zur Verbesserung des Umgangs mit der marktbeherrschenden Stellung der großen Informationsdienstleister sollen in diesem Teil I *Die Praxis des elektronischen Publizierens und ihre Probleme* in vier Beiträgen anhand folgender Fragen näher beleuchtet werden:

1. Welche Gefahr für die Datensouveränität der Wissenschaftler:innen geht vom ausufernden Datentracking der Wissenschaftsverlage aus und wie können Bibliotheken und andere Zugangsprovider damit umgehen?
2. Welche Gefahr liegt in den preisgetriebenen Zugangsbeschränkungen profitorientierter Wissenschaftsverlage und wie könnte die Verbreitung wissenschaftlicher Erkenntnisse durch eine öffentliche digitale Infrastruktur verbessert werden?

3. Wie kann durch Offenlegung der Publikationskosten die gewinnmaximierende Beliebigkeit der Preisfindung in Oligopolen verhindert werden?
4. Wie kann durch kooperatives Publizieren von Verlagen und Forschungsinstitutionen im Diamond-Open-Access die einseitige Abhängigkeit vermieden werden?

Im ersten Beitrag untersucht Eric Retzlaff die Unterschiede zwischen profitorientierter oder öffentlicher, digitaler Infrastruktur zur Verbreitung wissenschaftlicher Erkenntnisse. Während Wissenschaftsverlage durch den Impact Faktor den Wissenschaftler:innen Sichtbarkeit und Karrierechancen bieten, führt die Abhängigkeit von kommerziellen Anbietern zu erhöhten Zugangskosten – eine neue Gatekeeperfunktion, nun ausgeführt im Sinne kommerzieller Ziele und nicht mehr wissenschaftlicher (Einzel-) Interessen. Initiativen wie *ORCID, Zenodo* und *Open Research Europe* versuchen, öffentliche Strukturen bereitzustellen, stehen jedoch vor finanziellen und wettbewerblichen Hürden. Die Notwendigkeit zur Schaffung einer neuen, nicht-kommerziellen Infrastruktur wird dringlicher, je abhängiger die Wissenschaft von kommerziellen Interessen wird. Dies benötigt jedoch eine integrative Herangehensweise, welche die Stärken beider Infrastrukturen vereint, um eine gerechte, umfassende und möglichst nebenwirkungsfreie Wissensverbreitung zu gewährleisten.

Im zweiten Beitrag zeigt Stefan Drößler auf, weshalb es in letzter Zeit verstärkt Bemühungen um den Ausbau alternativer Publikationsinfrastrukturen zu den großen Verlagsoligopolen gibt. Dies ist eine Reaktion auf die unerwünschten Folgen der Transformationsverträge mit dem Verlagsoligopol: das hybride Publizieren wird nicht überwunden, eine Transformation vom Subskriptionsmodell zu Gold-Open-Access-Zeitschriften bleibt meist aus und die Erwerbungs- und Publikationsmittel werden weiterhin umverteilt zugunsten des Oligopols, so dass die Abhängigkeit von den Verlagen vertieft wird. Mit dem Co-Publishing von kommerziellem Verlag und öffentlicher Institution als strategischem Ansatz kann im sog. Diamond-Open-Access parallel ein alternativer Publikationspfad beschritten werden, wobei ein

Verhältnis von Koexistenz und Konkurrenz zwischen den beiden Pfaden zu erwarten ist.

Im dritten Beitrag diskutiert Bernhard Mittermaier die Problematik des Datentracking durch Wissenschaftsverlage. Welches sind die Motive der Verlage, die juristischen Aspekte und die bisherigen Reaktionen aus der Wissenschaft? Ergebnis der Betrachtung von insgesamt 14 Wissenschaftsverlagen in Nachfolge der 2023 abgeschlossenen DEAL-Verträge ist, dass in den meisten Fällen wesentlich mehr Daten gespeichert werden, als technisch und rechtlich notwendig ist, und dies auch exzessiv lang. Einige Verlage legen ganze Profile ihrer Nutzer:innen an. Die Datensouveränität der Wissenschaftler:innen ist dadurch ernstlich bedroht.

Im letzten Beitrag legt Lisa-Marie Stein das „Alphabet der Publikationskosten" offen. Mit dem vorgestellten Projekt „openCost: automatisierte, standardisierte Lieferung und offene Bereitstellung von Publikationskosten und Verlagsvereinbarungen" sollen Kostentransparenz und -kontrolle im wissenschaftlichen Publikationswesen erzielt werden. Bibliotheken, Publikationsbeauftragten und Wissenschaftler:innen kommt hierbei eine gestaltende Rolle zu: Sie können aktiv dazu beitragen, der Kommerzialisierung des wissenschaftlichen Publizierens entgegenzuwirken. Durch die Erfassung in einem standardisierten Metadatenschema und die technische Bereitstellung via OAI-Schnittstelle können Kostendaten frei zugänglich gemacht und automatisiert ausgetauscht werden. Die Abbildung der Kosten berücksichtigt dabei die Diversität der Wissenschaftsdisziplinen und ihrer medialen Repräsentationen. Trotz heterogener Publikationsobjekte, Workflows und Finanzstrukturen besteht ein hohes Maß der Konvergenz; Objekte und Muster können in einem gemeinsamen openCost-Standard beschrieben werden. Die strukturelle Ausgestaltung des Metadatenschemas und seine technische Integration erlauben inhaltliche Erweiterungen im dynamischen Umfeld der Open-Access-Transformation, FAIRe Prinzipien, aktuelle Daten sowie Informationsreichtum und die Reduzierung manueller Produktions- und Rechercheprozesse.

Die so eröffnete konstruktive Analyse der Infrastrukturen wissenschaftlichen Publizierens legt die informations- und soziotechnische Basis für die in Teil II betrachtete, davon abhängige Evaluierung wissenschaftlicher Expertise.

Literatur

Deutsche Forschungsgemeinschaft (DFG) (2021) „Datentracking in der Wissenschaft: Aggregation und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage. Ein Informationspapier des Ausschusses für Wissenschaftliche Bibliotheken und Informationssysteme der Deutschen Forschungsgemeinschaft. 28 Oktober 2021", online unter: https://www.dfg.de/download/pdf/foerderung/programme/lis/datentracking_papier_de.pdf.

Elsevier. https://www.elsevier.com/de-de/about vom 22.04.2022.

Holzer, Angela (2022) „Die Vermessbarkeit der Wissenschaft – Digitalisierung, wissenschaftliches Publizieren, Verhaltenstracking und Wissenschaftsbewertung", in: Mößner, N.; Erlach, K. (Hg.): Kalibrierung der Wissenschaft. Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis. Bielefeld, S. 163–181. DOI: https://doi.org/10.14361/9783839462102-007.

Retzlaff, Eric (2022) „Wer bewertet mit welchen Interessen wissenschaftliche Publikationen? Eine Skizzierung des Einflusses kommerzieller Interessen auf die Forschungsoutput-Bewertung", in: Mößner, N.; Erlach, K. (Hg.), Kalibrierung der Wissenschaft. Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis. Bielefeld, S. 139–162. DOI: https://doi.org/10.14361/9783839462102-006.

KAPITEL 2

Verbreitung wissenschaftlicher Erkenntnisse durch profitorientierte oder öffentliche digitale Infrastruktur?

Eric Retzlaff

1. Monopole, Abhängigkeiten und die Freiheit der Forschung

In der wissenschaftlichen Gemeinschaft wird die profitorientierte Wissenschaftsinfrastruktur durch Wissenschaftsverlage oft als zwiespältig angesehen. Renommierte Wissenschaftler/innen profitieren durch erhöhte Sichtbarkeit und Reichweite ihrer Arbeiten durch den immer noch vorherrschenden Impact Factor. Es ist sozusagen die Währung der Wissenschaft, um wissenschaftliche Expertise darzustellen und die wissenschaftliche Auswirkung auf das Wissenschaftsgebiet quantitativ zu messen (ZB Med 2025).

Die etablierte Form des wissenschaftlichen Publizierens wird oft als altbewährt angesehen und bietet den Forschenden eine

Zitationsvorschlag für dieses Kapitel:
Retzlaff, E. 2026. Verbreitung wissenschaftlicher Erkenntnisse durch profitorientierte oder öffentliche digitale Infrastruktur? In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* S. 23–42. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.c. License: CC BY 4.0

strukturierte Möglichkeit, ihre Karriere voranzutreiben (DFG 2022). Editoren, oft als „Door-Keeper" zum wissenschaftlichen System betrachtet, tragen eine große Verantwortung, da sie entscheiden, welche Forschungsleistungen wo und wie publiziert werden und somit sichtbar und anerkannt sind.

Jedoch ergibt sich aus den aktuellen Entwicklungen eine Marktkonzentration großer Verlage bis hin zu klaren Monopolen (Larivière 2015). Verlagsmonopole im wissenschaftlichen Bereich stehen in der Kritik, die sich in verschiedenen Befürchtungen und Ängsten widerspiegelt. Diese Sorgen entstehen durch die zunehmende Konzentration des Marktes auf wenige große Verlage, die dadurch eine dominante Stellung einnehmen.

In der aktuellen Diskussion über die Herausforderungen der wissenschaftlichen Publikationskultur stehen daher mehrere zentrale Ängste im Vordergrund. Zunächst besteht die Angst vor dem Verlust der Steuerung, da Wissenschaftler/innen und akademische Institutionen befürchten, die Kontrolle über ihre Forschungsergebnisse sowie deren Verbreitung zu verlieren. Wenn Verlage zu Gatekeepern des Wissens aufsteigen, könnte dies die Freiheit der Forschung und die Verbreitung neuer Erkenntnisse erheblich einschränken. Zwar ist der Einfluss der Wissenschaft über das Peer-Review-Verfahren und die Editoren-Tätigkeiten (noch) groß, jedoch entwickelt sich die geforderte Schnelligkeit des wissenschaftlichen Transfers mehr denn je zu einem reinen Geschäftsmodell, wo am Ende nicht die Qualität über die Veröffentlichung entscheidet, sondern die bloße Quantität. Das bedeutet im Umkehrschluss für die Wissenschaftsfreiheit, dass Forschungsergebnisse als reine Ware angesehen werden, die den Vermarktungsschienen der Großverlage dient und infolgedessen nicht unbedingt die Ziele des wissenschaftlichen Diskurses im Fokus stehen (Sandström 2016). Die Qualitätsorientierung sinkt dabei und die Verbreitung unter quantitativen Gesichtspunkten gewinnt an Bedeutung hinzu, wie das Beispiel MDPI Publishing zeigt (Fränti 2024). Dies könnte dazu führen, dass bestimmte Themen oder Ergebnisse bevorzugt werden, die primär den kommerziellen Interessen der Verlage dienlich sind und nicht primär dem

wissenschaftlichen Fortschritt. Eigentlich definiert sich die Rolle der Verlage aber in der „Veredelung" des Forschungsoutputs und wenn diese Rolle aus Geschäftsmodellgründen weniger wichtig wird, sinkt auch das Renommée der Wissenschaft insgesamt und ihre Ergebnisse erscheinen zunehmend beliebig bzw. interessengesteuert – aus gesellschaftlichen Gründen ist dies eine nicht zu unterschätzende Herausforderung (Cologna 2025).

Weiterhin besteht zweitens eine reale Angst vor Abhängigkeiten, die sich aus der Notwendigkeit ergibt, die Bedingungen großer Verlage zu akzeptieren. Diese Abhängigkeit kann die Kosten für den Zugang zu wissenschaftlichen Publikationen in die Höhe treiben und somit die Budgets von Bildungseinrichtungen stark belasten. Hierbei handelt es sich nicht nur um die bestehenden Paywall-and-Access-Modelle sondern auch um die Publikationshürden der Article-Processing-Charges (APCs) im Bereich des Open-Access-Publizierens. Steigen die Kosten weiterhin pro Jahr um den Faktor 5–10% (Butler 2024), wird es für wissenschaftliche Einrichtungen immer schwieriger, überhaupt Content bei den sogenannten renommierten Verlagen zu publizieren. Ausweichmodelle und eigene Initiativen könnten davon profitieren, werden allerdings nicht vom vorherrschenden Modell des Impact Factors berücksichtigt: Eigene, eventuell nicht kommerziell orientierte Infrastrukturen stehen in direkter Konkurrenz zu den profitorientierten (und im Markt vorherrschenden) Modellen. Damit würde man selektiv an der Finanzstärke einer Einrichtung die Qualität der Beiträge bemessen und eventuell disruptive oder ebenso relevante Beiträge ausschließen, weil es an finanziellen Möglichkeiten fehlt.

Ein weiteres zentrales Thema ist die Angst vor der Ausnutzung der skizzierten Marktmacht. Monopolistische Verlage könnten ihre Marktstellung nutzen, um Preise zu diktieren und ungünstig Bedingungen durchzusetzen, was zu einer Verzerrung des wissenschaftlichen Publikationsmarktes zum Nachteil kleinerer Akteure führen würde. Dies betrifft die Vielfalt des wissenschaftlichen Angebotes an sich, wenn man sich neue Themengruppen oder interdisziplinäre Forschungsthemen anschaut, die nicht

auf eine klar definierte Zielgruppe abzielen oder deutlich weniger Marktmacht besitzen, bspw. in Nischenthemen. Sind diese Angebote ab einem gewissen Grad wirtschaftlich relevant, bietet die Markstruktur oft keinerlei Abwehr gegen Übernahmen an. Beispiele aus der Praxis sind hier regelrechte Übernahmefeldzüge (Herb 2018) aus einem, wenn man es herunterbricht, stark steuerfinanzierten Geschäftsmodell: Die Marktmacht wird durch den Verkauf von Zugängen, Veröffentlichungsmöglichkeiten und Diensten getragen und das Geld dafür bringen im Großteil öffentlich finanzierte Einrichtungen auf, namentlich die Hochschulen und die außeruniversitären Einrichtungen wie Helmholtz-Gemeinschaft, Fraunhofer-Gesellschaft, Leibniz-Gemeinschaft oder Max-Planck-Gesellschaft. Die Sorge, dass Verlage zudem Daten über Forschungsaktivitäten und -präferenzen sammeln und diese Informationen für kommerzielle Zwecke nutzen könnten, verstärkt das allgemeine Misstrauen.

Kurzgefasst: Diese Ängste rühren von einem fundamentalen Konflikt zwischen dem Ziel der Wissenschaft, Wissen frei zu teilen, und den kommerziellen Interessen der Verlage, die ihr Geschäftsmodell auf dem Verkauf von Zugang zu diesem Wissen aufbauen. Dabei fahren sie zudem enorme Gewinnmargen ein (Walter, P. 2019), die letztlich dem Wissenschaftssystem entzogen werden.

2. Chancen und Herausforderungen nicht-kommerzieller Modelle

Die Frage, die sich hier stellt, ist natürlich: Könnten offene Strukturen ebenso Zugang zu Wissen bereitstellen, z.B. wenn sich die wissenschaftlichen Einrichtungen zusammenschließen oder gar gemeinnützige Plattformen aufbauen?

Die nicht-kommerziellen (oder not-for-profit-orientierten) wissenschaftlichen Infrastrukturanbieter ORCID (ORCID 5), Zenodo (Zenodo 7), Open Research Europe (Open Research Europe 4) und Open Journals Systems (Open Journal System 3), um ein paar

Beispiele zu nennen, stellen wichtige und gelungene Infrastruktur-Projekte dar, die versuchen, der dominierenden Präsenz profitorientierter Verlage entgegenzuwirken:

- ORCID bietet Forschern eine eindeutige Identifikationsnummer, die es ermöglicht, ihre Arbeiten und Beiträge effizient zu verknüpfen. Dies trägt zur Sichtbarkeit wissenschaftlicher Arbeiten bei und erleichtert die Zuordnung von Leistungen zu einzelnen Forschern. Selbst die Marktführer mit eigenen Identifikationssystemen, also Scopus mit Scopus-ID und Clarivate mit ResearcherID, konnten sich bisher nicht durchsetzen, waren fehleranfällig und unterstützen daher mittlerweile die ORCID-Initiative für ihre eigenen Produkte.
- Zenodo ist ein Open-Access-Repository, das darauf abzielt, Forschungsdaten und -ergebnisse zu speichern und zu teilen, wodurch der Zugang zu wissenschaftlichem Wissen verbessert wird. Es bietet für all diejenigen Einrichtungen einen schnellen Weg der Veröffentlichung an, die keine eigene OA-Repositories betreiben.
- Open Research Europe ist eine Plattform, die speziell für die Veröffentlichung von Open-Access-Forschungsartikeln in Europa entwickelt wurde, um den Wissenschaftlern eine transparente und zugängliche Publikationsmöglichkeit zu bieten. Die Plattform ist eng mit den eigenen Angeboten der EU-Forschungsförderung verknüpft. Open Research Europe ist z.B. Teil des OpenAIRE-Netzwerks. Das bedeutet, dass die Veröffentlichungen auf Open Research Europe in das OpenAIRE-Ökosystem integriert sind, was eine bessere Auffindbarkeit und Verlinkung zu anderen offenen Ressourcen und Daten ermöglicht.
- Und als letztes Beispiel sei das Open Journals Systems (OJS) für die Verwaltung und Veröffentlichung von Open-Access-Zeitschriften genannt. OJS spielt eine entscheidende Rolle in der wissenschaftlichen Veröffentlichungspraxis, insbesondere im Kontext der Open-Access-Bewegung. OJS bietet Institutionen und Forschern die Möglichkeit, ihre eigenen Zeitschriften unabhängig von kommerziellen Verlagen zu betreiben, was die akademische Autonomie stärkt und zu einem offenen, zugänglichen und transparenten Publikationssystem beiträgt.

Dies sind nur exemplarisch ein paar Beispiele, mit Sicherheit würde man viele weitere Projekte, Plattformen und Initiativen

finden; die Aufzählung ist nur ein kleiner Blick auf eine nicht-profitorientierte Lösungen in der Veröffentlichungspraxis.

Allerdings reicht dies nicht annähernd aus, das derzeitige System umzuwerfen und das hat zahlreiche Gründe. Denn trotz der Bedeutung dieser nicht-kommerziellen Anbieter stehen sie vor mehreren Herausforderungen, die mit der dominierenden Stellung kommerzieller Verlage im wissenschaftlichen Publikationswesen verbunden sind. Eine der größten Hürden ist die finanzielle Unterstützung. Viele nicht-kommerzielle Projekte haben Schwierigkeiten, ausreichende Mittel zu finden, um ihre Infrastruktur und Dienstleistungen aufrechtzuerhalten. Dies kann dazu führen, dass sie in ihrer Entwicklung und ihrem Wachstum eingeschränkt sind. Hat die Wissenschaft tatsächlich die Aufgabe, selbst Systeme zu entwickeln und aufrechtzuerhalten und kann sie es wirklich besser als kommerzielle Mitbewerber? Denn hier stellt sich die Frage: wer finanziert diese infrastrukturellen Angebote dauerhaft und trägt somit auch die Verantwortung und das finanzielle Risiko in der Fortführung der Angebote? Aus welchem Steuertopf müsste man diese Angebote weiterhin fördern, wenn ernstzunehmende Konkurrenzangebote entstehen sollen? Wer entscheidet darüber, welches Projekt oder welches Angebot wie gefördert wird? Können das der Staat oder die Wissenschaftsgemeinschaft leisten?

Als gut darstellbares Beispiel für die Beantwortung dieser Fragen eignet sich arXiv (arXiv 1). Die Plattform ist ein bekanntes Beispiel für eine nicht-kommerzielle wissenschaftliche Infrastruktur, die sich auf die Bereitstellung von Open-Access-Publikationen konzentriert. Die schon in den 80er-Jahren von der Cornell University errichtete Plattform ermöglicht Forschern, Preprints in verschiedenen Disziplinen, insbesondere in den Bereichen Physik, Mathematik und Informatik, hochzuladen und zu teilen. Das Finanzierungsmodell basiert dabei hauptsächlich auf einer Kombination aus institutioneller Unterstützung und Fördermitteln. Viele akademische Institutionen, darunter Universitäten und Forschungsinstitute auch aus Deutschland, wie bspw. wie Technische Informationsbibliothek Hannover (TIB), die die Finanzierung für die deutschen Hochschulen und weitere Einrichtungen

organisiert, tragen zur Finanzierung bei, indem sie Mitgliedsbeiträge zahlen. Diese Mitgliedsbeiträge ermöglichen es arXiv, die Betriebskosten zu decken, einschließlich Serverkosten, Wartung und Entwicklung. Zusätzlich erhält arXiv auch Fördermittel von verschiedenen Organisationen, die daran interessiert sind, die Open-Access-Bewegung und den freien Zugang zu wissenschaftlicher Literatur zu unterstützen. Diese Fördermittel helfen, die Unabhängigkeit und Nachhaltigkeit von arXiv zu sichern, sodass die Plattform nicht auf kommerzielle Einnahmen angewiesen ist. 2023 wurden $2.690.592 eingenommen, aber $3.359.075 ausgegeben, das Defizit Betrug also $668.483 (arXiv 2023). Auch für 2024 wird ein Minus-Betrag erwartet, der über weitere Spendenbeiträge verringert werden muss. Genau hier stellt sich aus finanzieller Sicht die Frage, wie weit die wissenschaftlichen Einrichtungen selbst Träger dieser Plattformen sein können oder müssen. Klar ist, Infrastruktur benötigen stabile finanzielle Rahmenbedingungen und Spielräume, diese sind aber natürlich nicht beliebig und vor allem finanziell beschränkt.

Ein weiteres Problem ist gar nicht mal finanzieller Art, sondern betrifft die Sichtbarkeit und Anerkennung der Plattform an sich. Nicht-kommerzielle Anbieter kämpfen oft darum, im Vergleich zu großen, profitablen Verlagen wahrgenommen zu werden. Diese Verlage verfügen über größere Ressourcen und ausgeklügelte Marketingstrategien, die es ihnen ermöglichen, ihre Produkte und Dienstleistungen besser zu positionieren. Der Wettbewerbsdruck, den sie ausüben, erschwert es den nicht-kommerziellen Anbietern, sich in einem Wettbewerbs-Markt zu behaupten. So finden sich sicherlich in einigen Nischen innovative Tools oder Produkte, die dann jedoch wiederum mit Angeboten kommerzieller Anbieter, bspw. ausgestattet mit hohem Risikokapital aus der Start-Up-Szene, konkurrieren. Das erschwert langfristig den Aufbau dieser Alternativen, wenn nicht finanziell stark unterstützt und politisch gewollt wird, wie dies bspw. bei den Angeboten der EU-Forschungsförderung der Fall ist. Insgesamt bleibt die Etablierung einer nachhaltigen nicht-kommerziellen Infrastruktur im wissenschaftlichen Bereich eine

komplexe Aufgabe, die durch finanzielle und wettbewerbliche Herausforderungen erschwert wird.

3. Die Rolle von CoARA und die Zukunft der wissenschaftlichen Exzellenzbeurteilung

Was jedoch den zentralen Punkt darstellt, ist die Zuschreibung des Renomées über etablierte Publikationsorgane und den eingangs genannten Impact Factor. Genau dieses System steht daher in der Kritik durch CoARA, die Coalition for Advancing Research Assessment (CoARA 2). Sie ist eine Initiative, die darauf abzielt, die Art und Weise, wie Forschungsleistungen bewertet werden, grundlegend zu verändern, um damit aus der Abhängigkeit etablierter Veröffentlichungsplattform zu kommen. Diese Koalition setzt sich aus einer Vielzahl von Akteuren des Wissenschaftssystems zusammen, darunter Forschungseinrichtungen und Förderorganisationen. CoARA verfolgt das Ziel, alternative Bewertungskriterien zu entwickeln (oder zumindest zu propagieren), die über den traditionellen Impact Factor hinausgehen, der weiterhin als Hauptmaßstab für die Qualität und den Einfluss wissenschaftlicher Arbeiten angesehen wird.

Die Wichtigkeit eines neuen, nicht impact-factor-orientierten Ansatzes ergibt sich aus mehreren zentralen Aspekten. Erstens fördert der Impact Factor häufig die Veröffentlichung in einer begrenzten Anzahl hochrangiger Zeitschriften, während er die Vielfalt der Forschungsdisziplinen und -ansätze ignoriert. Dies kann dazu führen, dass wertvolle Forschungsergebnisse aus weniger bekannten oder neueren Disziplinen nicht die Anerkennung erhalten, die sie verdienen (Larivière 2019). Ein neuer Ansatz würde demgegenüber eine breitere Palette von Forschungsleistungen wertschätzen und fördern. Die Fokussierung auf den Impact Factor führt zweitens oft zu einer Überbetonung quantitativer Ergebnisse, wobei qualitative Aspekte der Forschung vernachlässigt werden könnten (Sahel 2011). Dies kann dazu führen, dass Forscher sich gezwungen fühlen, ihre Arbeit auf quantifizierbare Erfolge auszurichten, anstatt innovative oder risikobehaftete

Forschungsansätze zu verfolgen. CoARA zielt daher darauf ab, die Qualität der Forschung angemessen (oder auch innovativer und gezielter) zu bewerten und somit einen Raum für kreative und transformative wissenschaftliche Arbeiten zu schaffen, die im jetzigen System wenig Berücksichtigung finden.

Nicht überall in der Wissenschaft ist das System der Exzellenzbeurteilung maßgeblich für den Erfolg verantwortlich. Für Infrastrukturanbieter hingegen spielt es eine wesentliche Rolle, um neue Autoren/innen für ihre Publikationsorgane zu gewinnen. Derzeit haben jedoch vor allem kommerzielle Anbieter wie Elsevier mit Scopus und Clarivate mit dem Web of Science marktbeherrschende Angebote, wenn es um die Erstellung der Impact-Faktoren geht. Diese Abhängigkeit von kommerziellen Anbietern trägt zur bestehenden Ungleichheit im Wissenschaftssystem bei und verstärkt den Druck auf Forscher, sich an den Kriterien der etablierten Bewertungssysteme zu orientieren.

Darüber hinaus spielt die Förderung von Open Science eine zentrale Rolle in den Bestrebungen von CoARA. Ein alternativer Bewertungsansatz könnte Open-Access-Publikationen und offene Forschungspraktiken stärker unterstützen, was zu einer breiteren Verteilung und Nutzung von Forschungsergebnissen führt. Dies ist besonders wichtig in einer Zeit, in der der Zugang zu wissenschaftlichem Wissen für eine breitere Öffentlichkeit von großer Bedeutung ist und KI-Tools genau auf diese Form von Forschungsergebnissen zugreifen. Das Problem: CoARA versucht, mehrere Hürden im Wissenschaftssystem, die durch die derzeitige Bewertungslandschaft entstanden sind, gleichzeitig zu nehmen, und stößt dabei an Grenzen. Eine der größten Herausforderungen besteht in der fehlenden Einheitlichkeit der Bewertungsstandards. Derzeit gibt es keine klaren und einheitlichen Richtlinien für die Bewertung von Forschungsleistungen, was zu Inkonsistenzen und Ungerechtigkeiten führen kann. Hier spielen Fachcommunities und die gute wissenschaftlichen Praxis eine wesentliche Rolle, was sich allerdings nicht immer in ein allgemeingütiges Schema packen lässt – ein großer Vorteil des Impact-Factors, der leicht quantifizierbar und berechenbar ist, solange man eine ernstzunehmende Datenbasis des wissenschaftlichen Outputs erhält. Und genau diese Datenbasis

wird durch Scopus oder das Web of Scinece seit Jahrzehnten zur Verfügung gestellt (Pranckutė 2021).

Insgesamt ist CoARA sicherlich eine wichtige Initiative auf Seiten der Wissenschaftscommunity selbst, die das Ziel verfolgt, das Wissenschaftssystem gerechter, transparenter und inklusiver zu gestalten. Durch die Förderung alternativer Bewertungsansätze, die über den traditionellen Impact Factor hinausgehen, könnte CoARA einen größeren Beitrag zur Weiterentwicklung der wissenschaftlichen Publikations- und Bewertungskultur leisten. Das Problem ist derzeit nur: Dieses System ist noch nicht in der Praxis etabliert oder gar anerkannt und stellt damit momentan noch keine funktionierende Alternative zum jetzigen System dar. Bildlich gesehen müsste man die Reifen eines Rennwagens während der Fahrt wechseln, damit das System Wissenschaft, in diesem Fall dargestellt durch den Rennwagen, nicht zum Stehen kommt.

Immerhin ist das Thema schon auf der politischen Agenda angekommen (European Commission 2024) mit dem Ziel, die Kriterien und Methoden zur Bewertung von Forschungsleistungen tatsächlich zu reformieren und um eine ausgewogenere und qualitativere Bewertung zu ermöglichen. Die bisherige starke Fokussierung auf quantitative Messungen wie Publikationszahlen und Zitationen wird dabei aktiv hinterfragt und soll durch qualitative Aspekte, die eine differenziertere Einschätzung der Forschungsarbeit erlauben, ergänzt werden. Als zentraler Punkt wird die Vielfalt der Bewertungskriterien betont, indem unterschiedliche Indikatoren und Kriterien für die Bewertung einbezogen werden. Man merkt also: Der Wille zur Veränderung ist vorhanden, jedoch sprechen wir hier über ein etabliertes, akzeptiertes System, das für Reportingaufgaben etabliert und auch für Karrierewege mitentscheidend ist und daher nicht so schnell und einfach ausgetauscht werden kann.

4. Vielfalt, Innovation und Unabhängigkeit im Fokus

In der aktuellen Diskussion über die Zukunft der wissenschaftlichen Infrastruktur lassen sich drei zentrale Thesen aus Sicht der

Wissenschaft formulieren. Erstens streben wir eine wissenschaftliche Landschaft an, die keine Oligopole zulassen sollte. Dies bedeutet, dass die Wettbewerbsbedingungen so gestaltet werden müssen, dass Vielfalt und Pluralität in der Forschung gefördert werden, um monopolartige kommerzielle Strukturen auszuschließen und das über Steuergelder finanzierte Geschäftsmodell dieser Strukturen (Larivière 2015) zu vermeiden.

Zweitens ist es entscheidend, innovative Infrastrukturen zu entwickeln, die die Anforderungen der zukünftigen wissenschaftlichen Herausforderungen erfüllen können. Diese Infrastrukturen sollten flexibel und anpassungsfähig sein, um den dynamischen Entwicklungen in der Wissenschaft gerecht zu werden und die notwendige Unterstützung für neue Forschungsansätze zu bieten.

Drittens ist es unerlässlich, dass wissenschaftliche Interessen Vorrang vor wirtschaftlichen Überlegungen haben, insbesondere wenn es um die Steuerung und Bewertung der Wissenschaft geht. Die Integrität und Unabhängigkeit der Forschung müssen gewahrt bleiben, damit wissenschaftliche Erkenntnisse möglichst unbeeinflusst von wirtschaftlichen Interessen der Verwertung entstehen können und „Wahrheiten" nicht als gekauft erscheinen.

In der Diskussion um die Finanzierung öffentlich getragener Infrastrukturen wird deutlich, dass diese im Vergleich zu kommerziellen Diensten mit erheblichen Herausforderungen konfrontiert sind: Sie sind in der Regel nicht so gestaltet, dass sie eigenständig auf Marktanpassungen reagieren können, sondern werden über klar formulierte Projekte oder relativ starre Vorgaben und Rahmenbedingungen in guter alter kameralistischer Tradition geplant. Kommerzielle Infrastrukturanbieter hingegen müssen per se flexibel und schnell auf die dynamischen Anforderungen des Marktes reagieren können, um ihre Gewinnziele erreichen zu können. Diese zielorientierte Agilität steht im Kontrast zu drei wesentlichen Nachteilen öffentlicher Infrastrukturen:

- Die Förderung von Services in öffentlichen Infrastrukturen erfolgt häufig schrittweise, was die Implementierung neuer Technologien und Dienstleistungen verlangsamt, da für jede Erweiterung erst

einmal Gelder akquiriert werden müssen. In der Regel fehlt es an Risikokapital und einer entsprechenden Risikobereitschaft, insbesondere wenn es um den Einsatz von Steuergeldern geht. Diese langsame Reaktionsfähigkeit kann dazu führen, dass öffentliche Infrastrukturen hinter den sich schnell entwickelnden Anforderungen zurückbleiben, veralten und dadurch uninteressant werden.

- Der Fachkräftemangel ist ein kritisches Problem. Es wird zunehmend schwierig, qualifiziertes Personal zu finden, insbesondere in spezialisierten Bereichen wie der Informatik. Das gilt verstärkt für befristete Stellen unter Berücksichtigung der eingeschränkten Möglichkeiten des öffentlichen Dienstes bei der Entlohnung.
- Oft besteht eine Herausforderung in einer gesicherten Folgefinanzierung. Während kommerzielle Dienste in der Lage sind, ihre Finanzierungsmodelle dynamisch anzupassen, fehlt der öffentlichen Infrastrukturen häufig eine langfristige Planungssicherheit. Dies führt zu Unsicherheiten, die die Entwicklung und Aufrechterhaltung innovativer Dienstleistungen behindern.

Zusätzlich ist zu beachten, dass öffentliche Infrastrukturen nicht marktorientiert arbeiten und daher möglicherweise nicht die notwendige Innovationskraft aufbringen, um mit den sich wandelnden Bedürfnissen der Wissenschaft Schritt zu halten. Um die Finanzierung öffentlich getragener Infrastrukturen zu sichern und ihre Wettbewerbsfähigkeit zu erhöhen, sind daher neue Ansätze erforderlich, die sowohl Risikokapital bereitstellen als auch Anreize für Innovation und Fachkräfteentwicklung schaffen. Hierbei könnte eine Kombination aus staatlicher Förderung, öffentlich-privaten Partnerschaften und einer strategischen Neuausrichtung der Finanzierungsmodelle hilfreich sein, um die Herausforderungen der Zukunft erfolgreich zu meistern.

5. Innovationskraft durch Risikokapital

Die Ausgründung von Google aus der Stanford University im Jahr 1998 stellt ein paradigmatisches Beispiel für die Entwicklung innovativer Technologien im Kontext der Risikokapitalfinanzierung dar. Ursprünglich als Forschungsprojekt von Larry Page und

Sergey Brin ins Leben gerufen, wurde Google schnell zu einem, wenn nicht sogar dem führenden Akteur im Bereich der Internetsuche. Der entscheidende Faktor für den Erfolg von Google war die frühzeitige und signifikante Unterstützung durch Risikokapitalgeber, die bereit waren, in das vielversprechende Geschäftsmodell zu investieren (Vise 2005). Diese Investitionen ermöglichten es Google sehr früh, seine technologischen Entwicklungen zügig voranzutreiben, die Infrastruktur schnell auszubauen und talentierte Fachkräfte einzustellen. Während Risikokapitalgeber bereit sind, hohe Risiken einzugehen, um potenziell hohe Renditen zu erzielen, müssen öffentliche Institutionen oft strenge Haushaltsvorgaben und politische Auflagen einhalten, die eine agile Reaktion auf neue Herausforderungen erschweren. Diese Unterschiede in der Finanzierung und den damit verbundenen Risikobereitschaften führen zu einer Kluft in der Innovationskraft zwischen privatwirtschaftlich und öffentlich finanzierten Infrastrukturen, was die Notwendigkeit unterstreicht, neue Ansätze zur Förderung von Innovationsprozessen in der öffentlichen Infrastruktur zu entwickeln.

Nun ist die Google-Erfolgsgeschichte vielleicht nicht das passende Beispiel für den Vergleich mit wissenschaftlicher Infrastruktur, sondern mehr ein Beispiel der Veränderung unserer Welt im Bereich der Digitalisierung. Aber genau hier lassen sich ebenso Beispiele finden, die unseren Forschungsalltag stark beeinflussen oder dies noch werden: Der Bereich der künstlichen Intelligenz ist prädestiniert für richtungsweisende Tools oder Unterstützungsplattformen im Bereich wissenschaftlichen Infrastrukturen, weil die Analyse großer Datenmengen (die Wissenschaft hat enormes, verborgenes Wissen) prädestiniert ist, maschinell durchgeführt zu werden (Lu 2024).

Hier sind es nicht unbedingt immer die größten Lösungen, die auf die Forschungslandschaft selbst einen signifikativen Einfluss haben. Am Beispiel Scite.ai (Scite 5a) lässt sich zeigen, wie eine innovative Plattform auf die Verbesserung der wissenschaftlichen Recherche und des Peer-Review-Prozesses abzielt. Die Hauptfunktion von Scite.ai besteht darin, die Zitationen

wissenschaftlicher Arbeiten zu analysieren und zu klassifizieren. Die Plattform geht über die herkömmliche Zitationsanalyse hinaus, indem sie nicht nur zählt, wie oft eine Studie zitiert wurde, sondern auch den Kontext der Zitationen untersucht. Scite.ai unterscheidet zwischen unterstützenden, gegensätzlichen und allgemeinen Zitationen, was Forschern ermöglicht, die wissenschaftliche Relevanz und den Einfluss von Arbeiten besser zu verstehen. Zusätzlich bietet die Plattform Funktionen wie die Erstellung von Literaturübersichten und die Identifizierung relevanter Forschungsthemen, was die Effizienz und Qualität der wissenschaftlichen Arbeit steigern kann.

Um diese innovative Technologie zu entwickeln und weiterzuentwickeln, war Scite.ai auf signifikante Investitionen durch Risikokapitalgeber angewiesen. Bisher hat das Unternehmen mehrere Finanzierungsrunden erfolgreich abgeschlossen und insgesamt über 4,2 Millionen US-Dollar Kapital akquiriert (Crunchbase 2a), fast die Hälfte ist dabei reines Risikokapital. Diese Mittel wurden verwendet, um die technischen Entwicklungen voranzutreiben, das Team zu erweitern sowie Marketing- und Vertriebsstrategien zu implementieren. Die Verfügbarkeit von Risikokapital ist für Unternehmen wie Scite.ai entscheidend, da sie es ihnen ermöglicht, in einem wettbewerbsintensiven Markt schnell zu wachsen und ihre innovativen Dienstleistungen weiter auszubauen. In einem Bereich, in dem öffentliche Fördermittel oft limitiert und an strenge Auflagen gebunden sind, verdeutlicht der Erfolg von Scite.ai, wie wichtig private Investitionen für die Förderung technologischer Innovationen in der Wissenschaft sind.

Auch das Beispiel Elicit geht einen ähnlichen Weg: die Plattform zielt darauf ab, den Prozess der wissenschaftlichen Recherche und das Sammeln von Evidenz zu optimieren. Die Plattform bietet Schnittstellen für systematische Literaturübersichten und extrahiert auf Knopfdruck relevante Informationen aus wissenschaftlichen Studien. Elicit ermöglicht es Nutzern, Forschungsfragen zu formulieren, relevante Studien zu identifizieren und die Ergebnisse in einer organisierten Weise zu analysieren. Durch

Funktionen wie automatisierte Datenextraktion und die Erstellung von Zusammenfassungen versucht Elicit Wissenschaftler/innen dabei zu unterstützen, Zeit zu sparen und die Qualität ihrer wissenschaftlichen Arbeiten zu verbessern. Auch Elicit ist von (viel) Risikokapital abhängig: In mehreren Finanzierungsrunden hat das Unternehmen insgesamt über 9 Millionen US-Dollar an Risikokapital gesammelt (Crunchbase 2b). Wären beide Beispiele auch über die traditionelle Projektförderung in der Wissenschaft möglich? Aus den oben genannten Gründen der Agilität in der unternehmerischen Projekt- und Risikoplanung (Elicit hat bspw. mehrfach seine Ausrichtung und Dienste geändert und angepasst) darf dies berechtigt angezweifelt werden.

Dennoch: Das alles wäre nicht möglich, wenn es keine freie Datenbasis gäbe, an der sich Beispiel-Plattformen wie Scite.ai oder Elicit bedienen können. Im Hintergrund greifen beide Dienste auf die Datenbasis von Semantic Scholar zu. Semantic Scholar ist eine umfassende, KI-gestützte akademische Suchmaschine, die Zugriff auf eine Vielzahl von wissenschaftlichen Artikeln und deren Zitationen bietet. Sie profitiert damit erheblich von der Open-Access-Bewegung und der freien Zugänglichkeit auf Repositorien, Preprint-Server wie arxiv.org oder weiteren Open-Access-Quellen. Semantic Scholar wird dabei hauptsächlich durch das „Allen Institute for AI (AI2)" finanziert, eine gemeinnützige Organisation, die von Paul Allen, einem Mitbegründer von Microsoft, ins Leben gerufen wurde. AI2 hat das Ziel, Fortschritte in der künstlichen Intelligenz voranzutreiben und innovative Projekte im Bereich der KI und der Wissenschaft zu unterstützen – die Plattformen scite.ai oder Elicit sind nur Beispiele und stehen exemplarisch für eine Vielzahl an neuen Tools, die um die Gunst der Wissenschaft buhlen.

Die Finanzierung von Semantic Scholar umfasst sowohl private Mittel von AI2 als auch zusätzliche Förderungen durch Forschungsstipendien und Partnerschaften mit akademischen Institutionen und anderen Organisationen. Da es sich um eine gemeinnützige Initiative handelt, liegt der Fokus auf der Bereitstellung eines Tools für die wissenschaftliche Gemeinschaft, ohne

primär auf Gewinnerzielung ausgerichtet zu sein. Was bedeutet dies für die Nutzung freier Infrastruktur zur Wissensverbreitung, wenn man den Open-Science-Gedanken konsequent weiterverfolgt und damit maßvoll Infrastrukturprojekte fördert?

Auf der positiven Seite ermöglicht die Verbreitung von Wissen über eine freie Infrastruktur einer breiten Öffentlichkeit den Zugang zu Informationen, die sonst möglicherweise nicht verfügbar wären. Dies fördert Bildung und Forschung in verschiedenen Bereichen und trägt zur allgemeinen Wissensgesellschaft bei. Darüber hinaus können kommerzielle Akteure, die auf öffentlich zugängliche Daten und Inhalte zugreifen, auf dieser Basis innovative Produkte und Dienstleistungen entwickeln, wie dies auch oben beschrieben ist. Diese Innovationen können wiederum zur Lösung gesellschaftlicher Herausforderungen beitragen. Zudem kann die Kommerzialisierung von Inhalten dazu beitragen, die Finanzierung von Forschungsprojekten und Infrastruktur zu sichern, da Einnahmen aus der Weiterverwertung zurückfließen und reinvestiert werden können, um die Qualität und den Umfang der bereitgestellten Inhalte zu verbessern. Dies ist für die Zusammenarbeit und Steuerung zukünftiger Infrastrukturen absolut relevant: Es sollte immer auch der Transfer als Teil des Vorhabens mitgedacht werden. Der Nutzen orientiert sich bei dieser Vorgehensweise an allen Dimensionen der Vierfach-Helix (Quadruple Helix) des Innovationsmodells von Schroth (Schroth 2021), also dem Zusammenwirken der vier Hauptakteure Wissenschaft, Industrie, Politik und Gesellschaft. Genau diese integrative Betrachtung würde deutlich mehr Anknüpfungspunkte zur finanziellen Ausgestaltung nachhaltiger Infrastrukturen bereitstellen.

Am Beispiel der Herausforderung des wissenschaftlichen Bewertungssystems müsste ein neues System, wie es von CoARA und der Politik gefordert wird, sowohl wirtschaftliche wie auch gesellschaftliche Aspekte mehr in den Mittelpunkt seiner Bestrebungen nehmen, um die Akzeptanz zu steigern. Die Wissenschaft liefert das Bewertungssystem und die Politik nutzt dieses zur Vergabe von Fördermitteln und der Steuerung des Wissenschaftssystems. Wirtschaftlich agierende Unternehmen könnten hier die

entsprechenden technischen Systeme und Innovationen beisteuern, die auf den zwischen Wissenschaft und Politik ausgearbeiteten Regeln basieren. Am Ende profitiert die Gesellschaft von einem Wissenschaftssystem, das qualitative Bewertungen fördert, Vielfalt in Forschungsansätzen anerkennt, transparente Bewertungsverfahren entwickelt, Open Science unterstützt und interdisziplinäre Zusammenarbeit sowie nachhaltige Finanzierungsmodelle fördert. Wie könnte man dieses Zielbild also besser umsetzen?

6. Freie Kommunikation, offene Daten und synergetische Ansätze für eine gerechte Wissensverbreitung

Um die Rolle der Wissenschaft zu stärken, spielen mehrere zentrale Gedanken eine wesentliche Rolle. Zunächst ist die Förderung freier Kommunikationsprotokolle entscheidend, da Forschungsoutput zwar in Vielzahl vorhanden ist, jedoch noch nicht vollständig im Sinne von Open Science zugänglich ist. Open Science ist keine Einzelveranstaltung der Wissenschaft, sondern sollte als Grundlage für weitere Bestrebungen dienen, Forschungsleistungen nachzunutzen – und dies eben auch unter Berücksichtigung kommerzieller Interessen. Des Weiteren ist die Implementierung von Trusted Computing unerlässlich, um offene Protokolle und Nachvollziehbarkeit zu gewährleisten; auch die Forschungsbewertung muss transparenten Kriterien genügen.

Ein weiterer wichtiger Aspekt ist die Stärkung einer offenen Forschungsdateninfrastruktur, um künftige Monopole über Forschungsdaten im Gegensatz zu denen bei Publikationen zu vermeiden, da diese Daten der Schlüssel zu unserem Erkenntnisgewinn sind. Hier hat die Wissenschaft derzeit klare Vorteile bei Verwendung und Nachnutzung, weil keine Abhängigkeit wie im kommerziellen Veröffentlichungssystem der Großverlage besteht. Die Daten sind bei den Institutionen gespeichert und werden dort bereitgestellt. Es fehlt jedoch weiterhin an geeigneten technischen Lösungen, um auf Basis der Forschungsdaten Geschäftsmodelle

mit der Wirtschaft zu etablieren, auch wenn die Nationale Forschungsdateninfrastruktur (NFDI) hier auf jeden Fall bereits den richtigen Weg eingeschlagen hat (Stahl 2024).

Schließlich sollte betont werden, dass der Staat nicht die Wissenschaft wie ein Unternehmen führen sollte; stattdessen sollte er fördernde Rahmenbedingungen schaffen, ohne die inhaltliche Steuerung der Wissenschaft zu übernehmen. Diese Überlegungen sind entscheidend, um die Wissenschaft als eine zentrale Säule der Gesellschaft zu stärken und ihre Funktion im Dienst der Allgemeinheit zu festigen. Für die Verbreitung wissenschaftlicher Erkenntnisse erfordert dies eine integrative Herangehensweise, die sowohl profitorientierte als auch öffentliche digitale Infrastrukturen berücksichtigt. Diese These basiert auf der Erkenntnis, dass jede dieser Strukturen ihre eigenen Vorteile bietet: Während profitorientierte Akteure oft Agilität und Innovationskraft mit sich bringen, bieten öffentliche Infrastrukturen Zugang und Transparenz, die für die Wissenschaft unverzichtbar sind. Um die Rolle der Wissenschaft in der Gesellschaft zu stärken, ist es daher unerlässlich, das Beste aus beiden Welten zu vereinen. Nur durch eine synergetische Verbindung dieser Ansätze kann eine umfassende und gerechte Verbreitung wissenschaftlicher Erkenntnisse sichergestellt werden, die sowohl der wissenschaftlichen Gemeinschaft als auch der Gesellschaft als Ganzes zugutekommt.

Anmerkung

Mit ChatGPT 4.1 via FhGenie wurde ein internes Tool zur Unterstützung der zur sprachlichen Verbesserung des Textes verwendet. Für die Ideenfindung und Inhalte des Textes wurde keinerlei KI-Hilfe verwendet.

Literatur

arXiv (2023) Annual Report 2023. https://info.arxiv.org/about/reports/2023_arXiv_annual_report.pdf.

Butler, L.-A., Hare, M., Schönfelder, N., Schares, E., Alperin, J. P., & Haustein, S. (2024) Open dataset of annual article processing charges

(APCs) of gold and hybrid journals published by Elsevier, Frontiers, MDPI, PLOS, Springer-Nature and Wiley 2019–2023 (V1). Harvard Dataverse. https://doi.org/10.7910/DVN/CR1MMV.

Cologna, V., Mede, N. G., & Berger, S. (2025) „Trust in scientists and their role in society across 68 countries", in: Nature Human Behaviour 9, S. 713–730. https://doi.org/10.1038/s41562-024-02090-5.

DFG (2022) Zur Gestaltung des wissenschaftlichen Publizierens – Positionspapier des Hauptausschusses der DFG. https://www.dfg.de/resource/blob/175770/5772e3980d4e81991dc716f94fbc2382/positionspapier-publikationswesen-data.pdf.

European Commission (2024) Report on research assessment, Publications Office of the European Union. https://data.europa.eu/doi/10.2848/931335.

Fränti, P. (2024) What is wrong with MDPI: Is it a predator or a serious competitor? arXiv preprint. https://arxiv.org/abs/2411.08051.

Herb, U. (2018) „Zwangsehen und Bastarde: Wohin steuert Big Data die Wissenschaft? Information", in: Wissenschaft & Praxis 69(2–3), S. 81–88. https://doi.org/10.1515/iwp-2018-0021.

Larivière, V.; Haustein, S.; Mongeon, P. (2015) „The oligopoly of academic publishers in the digital era", in: PLOS ONE 10(6), e0127502. https://doi.org/10.1371/journal.pone.0127502.

Larivière, V.; Sugimoto, C.R. (2019) „The Journal Impact Factor: A Brief History, Critique, and Discussion of Adverse Effects", in: Glänzel, W.; Moed, H.F.; Schmoch, U.; Thelwall, M. (Hg.): Springer Handbook of Science and Technology Indicators. Springer. https://doi.org/10.1007/978-3-030-02511-3_1.

Lu, C. et al. (2024) The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery. arXiv preprint. https://arxiv.org/abs/2408.06292.

Pranckutė, R. (2021) „Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world", in: Publications 9(1), S. 12. https://doi.org/10.3390/publications9010012.

Sahel, J.-A. (2011) „Quality versus quantity: Assessing individual research performance", in: Science Translational Medicine 3(84). https://doi.org/10.1126/scitranslmed.3002249.

Sandström, U.; van den Besselaar, P. (2016) „Quantity and/or quality? The importance of publishing many papers", in: PLOS ONE 11(11), e0166149. https://doi.org/10.1371/journal.pone.0166149.

Schroth, F.; Kaiser, S.; Schraudner, M. (2021) Quadruple Helix Collaboration in practice. Insights from the EU Project RiConfigure. Fraunhofer IAO. https://doi.org/10.24406/IAO-N-636131.

Stahl, F.; Hamann, A.; Hoff, K.; Kockelmann, K. (2024) Collaboration Models Between Industry and Academia. Whitepaper on behalf of the Section Industry Engagement of the National Research Data Infrastructure (NFDI). *Zenodo*. https://doi.org/10.5281/zenodo.10473579.
Vise, D.A.; Malseed, M. (2005) The google story. Random House. https://books.google.de/books?id=zyTCAlFPjgYC.
Walter, P.; Mullins, D. (2019) From symbiont to parasite: The evolution of for-profit science publishing. *Molecular Biology of the Cell* 30(20), S. 2537–2542. https://doi.org/10.1091/mbc.E19-03-0147.
ZB MED (2025) Journal Impact Factor und Alternativen. PUBLISSO. https://www.publisso.de/open-access-beraten/faqs/journal-impact-factor-und-alternativen.

Internetquellen

arXiv 1. arXiv. Verfügbar unter: https://arxiv.org.
CoARA 2. CoARA – Coalition for Advancing Research Assessment. https://coara.eu/.
Crunchbase 2a. Crunchbase Scite. https://www.crunchbase.com/organization/scite/financial_details.
Crunchbase 2b. Crunchbase Elicit. https://www.crunchbase.com/organization/elicit-52a6/financial_details.
Open Journal Systems 3. Open Journal Systems. https://pkp.sfu.ca/ojs/.
Open Research Europe 4. Open Research Europe. https://open-research-europe.ec.europa.eu/.
ORCID 5. Open Researcher and Contributor ID. Connecting Research and Researchers. https://orcid.org.
Scite 5a. Scite. AI for Research. https://scite.ai.
Semnatic Scholar 6. https://allenai.org/ai-for-science.
Zenodo 7. Zenodo. https://zenodo.org.

KAPITEL 3

Diamond Open Access?

Stefan Drößler

Die Wissenschaftskommunikation wird durch komplexe Aushandlungsprozesse um das elektronische Publizieren und den freien Zugang zu Veröffentlichungen (Open Access) auf spezifische Weise verändert. Dabei hat der Begriff "Diamond Open Access" in den vergangenen Jahren eine erstaunliche Karriere gemacht. Erstaunlich, weil mit dem Diamond-Ansatz eine radikale Infragestellung bisheriger Eigentumsverhältnisse für wissenschaftliche Fachzeitschriften einhergeht. Europäische Wissenschaftspolitik und Förderorganisationen haben die Analysen und Forderungen aus der Open-Access-Community aufgegriffen und entsprechende Programme auf den Weg gebracht. In diesem Umfang hat es wohl lange keinen so starken Rückenwind für ein nicht-kommerziell organisiertes wissenschaftliches Publikationssystem mehr gegeben, zumindest nicht in Deutschland und in der Europäischen Union.

Viele Analysen von Problemen des wissenschaftlichen Publikationssystems enden mit einer gewissen Ratlosigkeit und der

Zitationsvorschlag für dieses Kapitel:
Drößler, S. 2026. Diamond Open Access? In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum*. S. 43–68. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.d. License: CC BY 4.0

Einsicht in die Notwendigkeit, das Oligopol der Wissenschaftsverlage zu überwinden und Diamond Open Access (DOA) als Alternative auszubauen. Im Folgenden sollen zunächst zentrale Probleme aufgezeigt werden, um die es in der Open-Access-Diskussion geht (1). Anschließend werden konkrete Lösungsansätze und Hoffnungen, die mit DOA verbunden werden, beschrieben. Dazu gehört eine Debatte, was eigentlich unter dem Begriff zu verstehen ist (2). Schließlich wird der Frage nachgegangen, inwieweit mit DOA das Publikationssystem in Richtung einer aus Wissenschaftsperspektive wünschenswerten, von kommerziellen Zwängen und Logiken befreiten Kommunikationssphäre weiterentwickelt werden kann (3). Dafür wird an eine organisationstheoretische Analyse von Leonard Dobusch und Maximilian Heimstädt angeknüpft (Dobusch and Heimstädt, 2021), um zwei verschiedene Pfade des Open Access gegenüberzustellen: der kommerzielle, wie er in Europa und Nordamerika prägend ist und der nichtkommerzielle im Globalen Süden, bzw. konkreter in Lateinamerika. Daraus wird die These abgeleitet, dass die bisherigen Bemühungen um eine Systemtransformation des wissenschaftlichen Publikationssystems nicht zur Überwindung des kommerziellen Pfades, aber auf einen verstärkten Wettbewerb und Koexistenz beider Pfade hinauslaufen (4).

1. Enttäuschte Transformationsbemühungen

Niels Taubert unterscheidet vier „Wellen und Moden" der Open-Access-Transformation (vgl. Taubert 2024). Demnach ist die erste Welle geprägt von der Gründung institutioneller Repositorien, die zweite von Gold Open Access und Article Processing Charges (APC), also kommerzieller Angebote und dem Aufbau von Publikationsfonds, während die dritte Welle von Transformationsverträgen wie DEAL geprägt ist. Schließlich stellt er mit DOA gegenwärtig eine vierte Welle von Transformationshoffnungen fest. Dabei überlappen sich die Wellen in zeitlicher Abfolge. Kennzeichnend für jede Open-Access-Welle oder -Phase sind nach Taubert wiederkehrende Muster von hohen Erwartungen

und – am Ende jeder Phase – Enttäuschungen über die verfehlten Transformationsziele. Das Aufkommen der Diskussion um DOA lässt sich so verstehen als eine Reaktion auf die enttäuschten Hoffnungen, die mit den Transformationsverträgen in der dritten Phase verbunden waren. Dies soll im Folgenden inhaltlich näher bestimmt werden.

1.1 Gold Open Access vs. Hybrid Open Access

Ein zentrales Förderkriterium für die meisten Open-Access-Fonds an den Wissenschaftseinrichtungen ist bis heute der Ausschluss der Förderung von Publikationen in Hybrid-Zeitschriften. Das sind Zeitschriften, die nicht komplett auf Open Access umgestellt sind (Gold-Open-Access-Zeitschriften), sondern gegen eine Artikelgebühr Open Access anbieten, ansonsten aber das Closed-Access-Modell mit dem Zeitschriftenabonnement beibehalten. Dieses Geschäftsmodell, kritisiert als "Double Dipping" (vgl. Cressey 2009; Mittermaier 2015; Prosser 2015; Björk 2017), wird von den finanzierenden Bibliotheken in der Regel ausgeschlossen, nicht zuletzt auch deshalb, weil die DFG in ihren Förderkriterien Hybridzeitschriften ebenfalls ausschließt.[1] Eine Ausnahme bilden Hybrid-Zeitschriften, die Teil von Transformationsverträgen sind. Eine Hoffnung der Transformationsverträge, allen voran der DEAL-Verträge, war, dass die Wissenschaftsverlage diese Zeitschriften in den Gold-Open-Access "flippen". Diese Transformation ist bei DEAL bisher jedoch weitestgehend ausgeblieben. Von einer "Sackgasse" ist die Rede (vgl. Šimukovič 2024; Mittermaier 2025). Das derzeitige System halte die Wissenschaft in einem hybriden Modell gefangen, stärke die Marktmacht traditioneller Verlage, schränke den Wettbewerb ein, erhöhe die Eintrittsbarriere für den Open Access und erhöhe die Kosten für Bibliotheken und Universitäten (vgl. Pampel et al. 2024). Der Weg zum Gold Open Access bliebe, so die Einschätzungen, auf dem bisher eingeschlagenen Pfad absehbar unvollständig. Wilhelm Widmark, Vice-Chairman des Bibsam-Konsortiums, ein Zusammenschluss von 96 schwedischen Universitäten und Forschungseinrichtungen

unter dem Dach der schwedischen Nationalbibliothek, geht davon aus, dass viele Verlage bei den Transformationsverträgen am Mischsystem zwischen Closed und Open Access festhalten wollen (vgl. Schwamm 2023). Die Kraft zur Ablösung des Subskriptionsmodells durch die Aushandlung von verbesserten Transformationsverträgen ist offensichtlich erschöpft. Damit wird der „Transformationsvertrag" zu einem performativen Widerspruch und hat sich möglicherweise bis auf Ausnahmen („echte Transformationsverträge") überlebt.

Ein deutliches Warnzeichen für ein Scheitern der Transformationsverträge stellt auch das Beenden ihrer Unterstützung durch die Coalition S zum Jahresende 2024 dar (Plan S 2024). Schließlich war Robert-Jan Smits, früher europäischer Open-Access-Beauftragter und Plan-S-Begründer, mit dem Ziel angetreten, das hybride Publizieren zu „verbieten" (vgl. Else 2018). Die Hoffnung auf Veränderung war groß, wurde aber am Ende enttäuscht.

Konsequenterweise hat die Helmholtz-Gemeinschaft ihre Open-Access-Positionen angepasst und ein Memorandum zur Open-Access-Transformation veröffentlicht, um zu verbindlicheren Festlegungen für das Journal Flipping zu kommen:

„Die Helmholtz-Gemeinschaft wird ab dem Jahr 2026 Transformationsverträge nur dann abschließen bzw. ihnen beitreten, wenn sie belastbare Festlegungen zum Flipping enthalten. Ab dem Jahr 2028 wird sie Transformationsverträge nur dann abschließen bzw. ihnen beitreten, wenn sie den vollständigen Flip spätestens zum Vertragsende vorsehen." (Gemeinsame Task Group Open-Access-Transformation der Arbeitskreise Open Science und Bibliotheks- und Informationsmanagement und dem Helmholtz Open Science Office 2024: 4)

Das Bibsam-Konsortium hat seine Policy ebenfalls angepasst und will ab 2026 keine Read & Publish-Verträge mehr schließen, die Hybrid-Zeitschriften enthalten. Dagegen sollen alternative Geschäftsmodelle unterstützt werden, genauso wie die Migration von „researcher-driven journals" in traditionellen Verlagen zu anderen Plattformen (vgl. National Library of Sweden 2024).

1.2 Dekonstruktion von Transformationsmythen

Neben der ungelösten Problematik des Hybrid Open Access werden an Transformationsverträgen eine Reihe weiterer Schwierigkeiten diskutiert. Farley et al. machen „sechs Mythen" aus: die Vorstellung, dass die Transformationsverträge ein bewährtes Mittel für den Übergang von Closed zu Open Access sei, also das Hybrid-Thema, wird als ein Mythos identifiziert (1). Des Weiteren kritisieren sie die Vorstellung, dass Transformationsverträge zu einem gerechten wissenschaftlichen Publikationsökosystem führen (2) und das APC-Modell ablösen würden (3). Auch größere Kostentransparenz (4), mehr Preiswettbewerb (5) und die Vorstellung, dass Bibliotheken in bessere Verhandlungspositionen kommen würden (6), werden als Mythos verworfen (vgl. Farley et al. 2021: 298).

Diese sechs „Mythen" stecken ein Themenfeld ab, das über die Diskussion um die Transformationsverträge hinausweist und auch die bereits schon länger bestehenden systemischen Probleme des kommerziellen Open Access betreffen. Deren Überwindung wird von einem nicht ganz neuen Ansatz erhofft: dem Diamond Open Access.

2. Diamond Open Access als Lösungsansatz

2.1 Einführende Überlegungen

Der Begriff „Diamond Open Access" ist von Christian Fuchs und Marisol Sandoval eingeführt worden (vgl. Fuchs and Sandoval 2013), um dem profitorientierten wissenschaftlichen Publikationssystem mit „Monopolpreisen" und den ungleichen Zugangsmöglichkeiten zum Publizieren eine Alternative gegenüberzustellen. Sie halten den Begriff „Gold Open Access" für konzeptionell unzureichend, weil in einer Kategorie qualitativ unterschiedliche Modelle wie „for-profit" und „not-for-profit" vermischt und nicht weiter differenziert werden können. Fuchs und Sandoval kritisieren Open-Access-Vordenker

wie Peter Suber (ebd.: 438): „Gold Open Acess" solle besser in „Corporate Open Access" (COA) unbenannt werden und den „for-profit"-Teil von Gold bezeichnen. Künftig würde man nur noch zwischen COA, DOA und Green OA unterscheiden, wobei die Terminologie von Peter Subers Unterscheidung zwischen „gratis" und „libre" als weitere Dimension der Klassifizierung ergänzt werden könne. Sie definieren DOA wie folgt:

> „*In the Diamond Open Access Model, not-for-profit, non-commercial organizations, associations or networks publish material that is made available online in digital format, is free of charge for readers and authors and does not allow commercial and for-profit re-use.*" (ebd.: 438)

Printversionen sind nach dem Verständnis von Fuchs und Sandoval nicht ausgeschlossen, wenn nur die Druckkosten in Rechnung gestellt und keine Gewinne gemacht werden. Die digitale Publikation muss für Autor:innen, Bibliotheken und Leser:innen gebührenfrei bleiben. Die Nachnutzung kann über eine Creative-Commons-Lizenz geregelt werden, wobei eine nichtkommerzielle Nutzung sichergestellt werden muss. Die DOA-Definition von Fuchs und Sandoval ist damit auch bei der Nachnutzung ein „strictly not-for-profit modell" (ebd.: 438). Sie sehen zum Zeitpunkt ihrer Analyse 2013 DOA zwar als „dominante Realität" an, sehen aber die Gefahr, dass COA die Zukunft prägen wird. Deshalb brauche DOA politische Unterstützung: Finanzierungsmodelle, um Assistenten für Herausgeber, Lektoren, Korrektoren, Grafikdesigner und Techniker anzustellen sowie verpflichtende DOA Policies für die Finanzierung und Forschungsbewertung. Universitäten sollten in ihren Policies Herausgeber, Gutachter oder Tätigkeiten in Herausgebergremien für DOA besonders berücksichtigen. Die wichtige Rolle, die Zitationsdatenbanken spielen, sollte durch das Directory of Open Access Journals (DOAJ) abgelöst werden. Schließlich sollten weltweit Forschungsförderer DOA zur Pflicht machen (ebd.: 339 f.).

In normativer Hinsicht sind diese Ansätze, über das wissenschaftliche Publikationssystem im digitalen Zeitalter nachzudenken, nicht

neu. Insofern gibt es ein DOA „avant la lettre". So ist die „erste Welle des Open Access" (vgl. Taubert 2024) mit dem Aufbau der Repositorien (Publikationsserver) eng verbunden mit dem Verständnis einer öffentlich finanzierten Publikationsinfrastruktur, kostenfrei für Wissenschaftlerinnen und Wissenschaftler und unabhängig von Verlagen. So verfolgte auch Paul Ginsparg mit der Gründung des Preprint-Servers arxiv das Ziel, den Austausch von Informationen zu „demokratisieren" (Ginsparg 2011: 4), zunächst natürlich innerhalb der eigenen Fachcommunity.

2.2 Bestehende Services

Die Autorinnen und Autoren der „OA Diamond Journals Study" kommen zu der Einschätzung, dass weltweit zwischen 17.000 und 29.000 DOA-Zeitschriften verbreitet sind (Bosman et al. 2021: 27). In den Geistes- und Sozialwissenschaften ist das DOA-Modell bei den Fachzeitschriften im Vergleich zu Naturwissenschaften und Medizin am dominantesten. Die Zeitschriften werden oft von Universitäten und Fachgesellschaften verantwortet, während „big science"- und Medizin-Zeitschriften kommerziellen Verlagen gehören (ebd.: 33f). Westeuropa und Afrika haben den größten Anteil an APC-basierten Zeitschriften (je rund 45 Prozent), gefolgt von den USA und Kanada (zusammen 37 Prozent), während der Durchschnitt weltweit bei 27 Prozent APC-basierter Zeitschriften liegt. Die restlichen 73 Prozent sind DOA-Zeitschriften. Lateinamerika sticht hier mit 95 Prozent DOA-Anteil besonders hervor, sodass man bei der OA-Entwicklung von einem eigenen, unabhängigen DOA-Pfad sprechen kann.

1997 wurde in Brasilien SciELO (Scientific Electronic Library Online) gegründet.[2] Die Plattform wird vom Nationalen Rat für Wissenschaftliche und Technologische Entwicklung gefördert. DOA-Zeitschriften sind der Standard-Weg für die Veröffentlichungen.[3] Heute gehören 16 Länder zum SciELO-Netzwerk, über Lateinamerika hinaus auch Südafrika, Spanien und Portugal.[4] In Mexiko entstand 2003 mit Redalyc (**Red** de Revistas Científicas de **A**mérica **L**atina **y** El **C**aribe, España y Portugal) eine

ganz ähnliche digitale Bibliothek, die exklusiv jene Zeitschriften umfasst (Indexierung und Hosting-Service), die das Nonprofit-Publikationsmodell und Open Access teilen.[5] 2018 wurde entschieden, das Angebot über Lateinamerika, Spanien und Portugal hinaus auszuweiten auf alle Zeitschriften, die den Redalyc-Kriterien entsprechen, insbesondere DOA (Aguado-López et al. 2024: 2). Neben Redalyc kam 2018 AmeliCA hinzu. Das Ziel ist, wissenschaftliche Erkenntnisse als Gemeingut zu bewahren durch Stärkung eines offenen Publikationssystems für die gleichberechtigte Teilnahme am globalen wissenschaftlichen Dialog. Der Name AmeliCA leitet sich vom Nahuatl-Begriff „ameli" ab, der sich auf das Entstehen, insbesondere auf natürliche Wasserquellen, bezieht, CA steht für den spanischen Begriff für Open Science („Ciencia abierta").[6] Einen Überblick über weitere OA-Initiativen im Globalen Süden gibt Monica Berger (Berger 2021: 396–398).

In Deutschland wurde in letzter Zeit das DOA-Modell zunehmend unterstützt. Im Rahmen der Digital Peer Publishing Initiative des Landes Nordrhein-Westfalen (DiPP) wurden seit 2004 DOA-Fachzeitschriften gefördert (Nötzelmann 2007), mittlerweile abgelöst durch einen ähnlichen Service des Hochschulbibliothekszentrums (hbz).[7] Die Deutsche Forschungsgemeinschaft förderte nicht nur den Aufbau der Open-Access-Repositorien, sondern seit 2006 auch den Aufbau von Publikationsinfrastrukturen für Zeitschriften und Bücher, die man heute als DOA-Infrastruktur verstehen würde.

3. Neubelebung des Diamond-Ansatzes

3.1 Internationale wissenschaftspolitische Debatten und Initiativen zur Priorisierung von DOA

Auf unterschiedlichen Ebenen, global, europäisch, national, kam es in den vergangenen Jahren zu wissenschaftspolitischen Bekenntnissen, Aufforderungen und Aktivitäten in Richtung eines nichtkommerziellen Publikationssystems. Die UNESCO

initiierte 2024 die „Global Diamond Open Access Alliance", um DOA zu fördern.[8] Die Organisation der Vereinten Nationen für Bildung, Wissenschaft und Kultur (UNSECO) verbindet DOA mit den Werten von Open Science und den UN-Zielen für nachhaltige Entwicklung (Sustainable Development Goals, SDGs) als Teil des Völkerrechts. Zu den drei Schlüsselcharakteristiken gehören erstens der Community-Ansatz („community-led") und die akademische Verortung („academic-centric"), zweitens ein multilingual und kulturell inklusives Verständnis, um Bibliodiversität über Sprach- und Kulturgrenzen hinaus zu fördern sowie drittens Gerechtigkeit und Zugänglichkeit, um Forschenden aus Regionen mit wenig Ressourcen eine Teilhabe zu ermöglichen und damit die SDGs zu unterstützen. Mit der völkerrechtlichen Verankerung von DOA gehen für die Mitgliedsstaaten der UNESCO Berichtspflichten einher, um Fortschritte zu dokumentieren.

In den früheren Empfehlungen der UNESCO zu Open Science wird zwar nicht auf den Terminus „Diamond Open Access" Bezug genommen, die Mitgliedsstaaten werden jedoch aufgerufen, nicht-kommerzielle Open-Science-Infrastrukturen zu unterstützen und in diesem Bereich entsprechend zu investieren (UNESCO 2021: 23). Es folgte ein Jahr später der „Action Plan for Diamond Open Access", eine Initiative europäischer Forschungsförderorganisationen (vgl. Ancion et al. 2022). Im Action Plan werden Ziele in vier Themenfeldern benannt: Effizienz, Qualitätsstandards, Capacity building (Aufbau von Management-Fähigkeiten, Kommunikationsstrategien, Tools, ...) und Nachhaltigkeit. In dem Zusammenhang wurde die Einrichtung eines „Capacity Centre for Diamond Publishing (CCDP)" angekündigt. Nach der „OA Diamond Journals Study" (vgl. Bosman et al. 2021) kam die Idee auf, über einen „European Diamond Capacity Hub (EDCH)" die Grundlage für eine föderierte DOA-Infrastruktur zu schaffen. Dazu gehört die Stärkung der europäischen DOA-Community, die Unterstützung nationaler Capacity Hubs, Diamond-Verlage und Dienstleister.[9] Die EU-Projekte DIAMAS (Developing Institutional Open Access Publishing Models to Advance Scholarly Communication)[10] und

CRAFT-OA (Creating a Robust Accessible Federated Technology for Open Access) dienen der Umsetzung dieser Ziele.[11]

Die Regierungen der Europäischen Union betonen in ihren Ratsschlussfolgerungen 2023 die „Bedeutung gemeinnütziger Open-Access-Modelle des wissenschaftlichen Publizierens, bei denen Autorinnen und Autoren oder Leserinnen und Lesern keine Gebühren in Rechnung gestellt werden" (Rat der Europäischen Union 2023: 4). Die Mitgliedsstaaten werden ermutigt, die Publikationsplattform Open Research Europe (ORE) zu nutzen (ebd.: 7).

Die Hinwendung zu DOA auf europäischer Ebene ist auch Ergebnis der Kritik an Plan S, wie sie die mexikanische Wissenschaftlerin Arianna Becerril-García aus Sicht des Globalen Südens schon in der Planungsphase von Plan S formuliert hat. Sie befürchtete eine zunehmende Kommerzialisierung von OA (Smits and Pells 2022: 124f). Gold OA und Transformationsverträge sieht sie als Geschäftsmodelle, die die Wissenschaft privatisieren und Ländern des Globalen Südens schaden, während DOA Wissenschaft als öffentliches Gut ermöglicht. So sollten DOA-Zeitschriften auch als öffentliche Güter verstanden werden, frei von Gebühren für das Publizieren oder Lesen und hauptsächlich durch öffentliche Einrichtungen finanziert (Becerril García 2024: 455).

Der Globale Süden verschafft sich über große internationale Konferenzen zunehmend Gehör. Der erste Globale DOA-Gipfel 2023 in Mexiko versteht DOA als Rückbesinnung auf die ursprünglichen Ziele der Openness, die durch kommerzielle Interessen abgefälscht worden sind (vgl. Manifesto on Science as Global Public Good 2023). Beim zweiten DOA-Gipfel 2024 in Südafrika findet sich in der „Toluca – Cape Town Declaration on Diamond Open Access" die Auffassung, dass das Teilen von Wissen ein Menschenrecht ist. In Übereinstimmung mit den UNESCO-Empfehlungen zu Open Science als „community-owned, community-led, and non-commercial" verstanden. Darüber hinaus wird von DOA erhofft, ein Treiber für Dekolonisierung und Entmarginalisierung zu sein. Ein wichtiger Aspekt der Deklaration ist der Hinweis auf die Unterstützung durch Systeme der Wissenschaftsbewertung, die für die Umsetzung von

DOA notwendig sind (vgl. Toluca – Cape Town Declaration on Diamond Open Access 2024).

3.2 Wissenschaftspolitische Debatte in Deutschland für DOA als komplementäres Ergänzungsmodell

Die wissenschaftspolitische Debatte in Deutschland ist von unterschiedlichen Akteuren geprägt, die zu einer etwas nüchterneren Einschätzung von DOA kommen und der stark normativ geprägten Diskussion nicht in diesem Maße folgen. DOA erscheint dadurch lediglich als ein mögliches Geschäftsmodell unter anderen.

So sieht erstens der Wissenschaftsrat noch kein „Idealmodell" des Open Access entwickelt (Wissenschaftsrat 2022: 8) und setzt daher etwas andere Akzente: Gold OA solle nicht mit Artikelgebühren gleichgesetzt werden. Eine analytische Trennung von „Zugangsregimen" und Geschäftsmodellen wird vorgeschlagen (ebd.: FN 4). Die verbreitete Minimaldefinition von DOA als „Bezeichnung für Publikationsorgane, die auf Publikationsgebühren bzw. Article Processing Charges (APC) verzichten" wird deshalb konsequenterweise ergänzt: „Äquivalent dazu ist die Kombination von Goldenem Weg mit einem medienbezogenen Geschäftsmodell" (ebd.: 104). Die medienbezogenen Finanzierungsmodelle würden dann die Kernprobleme der publikationsbezogenen Geschäftsmodelle vermeiden: linearer Anstieg der Gesamtkosten mit steigenden Publikationszahlen, der damit verbundene Anreiz der Umsatzsteigerung und die mangelnde Bereitschaft der Forschenden, aufgrund von Reputationseffekten auf alternative Publikationsorte umzusteigen (ebd.: 72f). Der Wissenschaftsrat begrüßt deshalb neue DOA-Publikationsorte, auch zur Erhaltung der Diversität der Angebote und „zur Vermeidung unerwünschter Konzentrationsprozesse" (ebd.: 73). Sie sollten im Rahmen der Projektförderung mitfinanziert werden (ebd.: 80). Kritisch gesehen werden Mängel beim technischen Support, der Skalierbarkeit und nachhaltigen Absicherung der Angebote über

eine Projektförderung hinaus. Neu in die DOA-Debatte bringt der Wissenschaftsrat einen Wettbewerbsgedanken ein: Die Ausschreibung von Publikationsdienstleistungen für bestimmte Zeiträume durch die Herausgeber. „Dadurch würde ein Wettbewerb um kostengünstige und qualitätsvolle Publikationsdienstleistungen unabhängig von dem Renommee eines Herausgebergremiums ermöglicht werden" (ebd.: 73f). Der Wissenschaftsrat nimmt damit eine im Vergleich zu radikaleren Nonprofit-Ansätzen des DOA moderatere Haltung ein, auch weil DOA nicht als allein favorisiertes Publikationsmodell gesehen wird.

In den zweitens von der Kultusministerkonferenz verabschiedeten Leitlinien zu „Open Access in Deutschland" wird auf diese Analyse Bezug genommen: „Bund und Länder teilen die Einschätzung des Wissenschaftsrates, dass Diamond-Open-Access-Modelle zur Diversität des Systems beitragen und in Konkurrenz zu gebührenfinanzierten Modellen treten können." (Bundesministerium für Bildung und Forschung (BMBF) 2023: 4)

Zuweilen wird stärker zwischen wissenschaftsgeleitetem („community-led", „scholar-led") Open Access und DOA unterschieden. So sieht viertens die Allianz der Wissenschaftsorganisationen für die strategische Weiterentwicklung des wissenschaftlichen Publikationswesens als einen von vier Themenbereichen die Etablierung von wissenschaftsgeleiteten Open-Access-Infrastrukturen (Alexander von Humboldt Foundation et al. 2025: 1). „Wissenschaftsgeleitete Weiterentwicklung" fungiert dabei auch als Leitprinzip, das besagt, dass die Publikationsorgane „unter der Hoheit oder zumindest Mitsprache von Forschenden oder wissenschaftlichen Einrichtungen" (ebd.: 4) betrieben werden. Dazu gehöre der Besitz der Titelrechte einer Zeitschrift oder Buchreihe (ebd.: 12). In diesem Zusammenhang sieht die Allianz die Förderung nicht gewinnorientierter OA-Modelle als komplementär zu den Angeboten kommerzieller Verlage, wobei der Begriff Diamond Open Access keine Aussage darüber treffe, ob ein Publikationsorgan von einer öffentlichen Einrichtung oder privatwirtschaftlich verlegt werde. Ebenso sei „keine Aussage über mögliche Gewinnabsichten" (ebd.) getroffen. Die Allianz beschränkt sich damit auf eine Minimaldefinition

von DOA, die Publikationsorgane bezeichnet, die keine Publikationsgebühren erheben. Der Begriff DOA rückt damit gegenüber der wissenschaftsgeleiteten Governance von Open Access in den Hintergrund. Die weiteren Leitprinzipien der Strategie – Offenheit von Publikationen und Publikationsinfrastrukturen, Nachhaltigkeit durch Kosteneffizienz und Kostentransparenz, Förderung von Bibliodiversität und digitale Souveränität – stehen damit eng in Verbindung.

Die Fokusgruppe „scholar-led.network" stellt fünftens eine enge Verbindung von DOA und dem wissenschaftsgeleiteten Publizieren her, was sich im Begriff „Scholar-led-Diamond-OA" niederschlägt. Die Gruppe vertritt Wissenschaftlerinnen und Wissenschaftler im deutschsprachigen Raum, die in „kollaborativer Eigenregie" gebührenfreie Zeitschriften und Bücher verlegen und von Großverlagen unabhängig sind. Sie spricht aus der Perspektive geistes- und sozialwissenschaftlicher Fachdisziplinen, sieht eine strukturelle Benachteiligung durch die Bevorzugung von Großverlagen, versteht Forschung als öffentliches Gut und fordert Bibliodiversität sowie eine nachhaltige Finanzierung für DOA ein (vgl. Fokusgruppe Scholar-Led.Network 2021).

Schließlich hat sechstens die Leopoldina-Arbeitsgruppe „Zukunft des wissenschaftlichen Publizierens" einen neuen Ansatz zur Finanzierung von DOA vorgelegt. Sie folgt ebenfalls dem Prinzip der wissenschaftsgeführten Publikationsplattformen, das mit DOA verbunden wird: Betrieb und Finanzierung der Zeitschriften müssen in der Hand wissenschaftlicher Institutionen bleiben, die dafür ein Budget erhalten. Die Finanzierung wissenschaftlicher Publikationen wird als öffentliche Aufgabe verstanden, weil Wissenschaft ein öffentliches Gut sei. Demzufolge müssten auch wissenschaftliche Publikationen wieder zu einem öffentlichen Gut werden (Tautz et al. 2025: 16). Das Budget wird nach der gleichen Logik begründet, wie die Forschungsförderung, die ebenfalls auf Antrag durch öffentliche Mittel erfolgt und auch regelmäßig begutachtet wird. Antragsberechtigt wären Fachgesellschaften, Akademien oder öffentliche Einrichtungen mit „wissenschaftlichem oder infrastrukturellem Auftrag". Das schließt wissenschaftliche

Bibliotheken als aktive Akteure für die Bereitstellung einer gemeinnützigen Publikationsinfrastruktur ein, die aber zusätzlich auch eine Rolle für „Archivierungsaufgaben" haben, die separat finanziert werden sollten (ebd.: 15). Die Umsetzung einzelner „Publikationskomponenten" kann nach dem Leopoldina-Modell durch eine Vergabe an kommerzielle Dienstleister nach wettbewerblichen Verfahren erfolgen, so wie es der Wissenschaftsrat ebenfalls vorschlägt. Als zentraler Ansprechpartner für das DOA-Modell wird künftig die im Aufbau befindliche Servicestelle für Diamond Open Access (SeDOA) gesehen (ebd.: 13).

3.3 Kritik am DOA-Verständnis

Die Diskussion um Diamond Open Access ist Ausdruck eines Selbstverständigungsprozesses der Open Access Community über die Ziele und strategische Weiterentwicklung des wissenschaftlichen Publikationssystems. Neben den vielen hier skizzierten wissenschaftspolitischen Initiativen werden auch grundsätzliche Probleme diskutiert und Definitionen in Frage gestellt. Martin Paul Eve sieht im Begriff Diamond Open Access einen „kategorialen Irrtum". Während Gold und Green Open Access auf die Bedingungen des Zugangs abheben, verlagsseitig oder über ein Repositorium, bezeichne DOA die Finanzierungsbedingungen, unter denen ein Verlag operiere. Der Begriff trage auch zu der falschen Gewissheit bei, Gold Open Access sei mit Artikelgebühren verbunden (vgl. Eve 2021). Eve, mit Caroline Edwards Gründer der DOA-Plattform Open Library of Humanities (OLH), kritisiert die Abgrenzung zu kommerziellen Angeboten („anti-commercialism") sowie die Do-it-yourself-Mentalität und hält „high quality" DOA auf der Grundlage eines Geschäftsmodells mit kommerzieller Unterstützung für möglich. Der Kern der Kritik von Eve besteht darin, dass die Forschungsförderer nur die einzelnen Publikationen fördern. Stattdessen sei eine dauerhaft finanzierte Publikationsinfrastruktur nötig. Er fragt, wie es sein könne, dass ein Großteil der Infrastruktur so

ungerechtfertigt unterfinanziert ist und auf ehrenamtlicher Arbeit basiere, während Unternehmen wie Elsevier eine Gewinnspanne von 37 Prozent bei einem Umsatz von mehreren Milliarden Dollar erzielten (ebd.).

Dellmann et al. kritisieren ebenfalls den Begriff DOA, die Gleichsetzung von Gold OA und Artikelgebühren und folgen der analytischen Trennung des Wissenschaftsrats von Zugangsregimen und Geschäftsmodellen. Ziel dieser Bemühungen ist die Besinnung auf „Gold Open Access" als ausreichende Kategorie. DOA wird als Unterkategorie von Gold OA verstanden und der „Diamond-Separatismus" des Action Plan for Diamond OA (vgl. Ancion et al. 2022) abgelehnt, da begrifflich gar nicht mehr auf Gold OA Bezug genommen werde (Dellmann et al. 2022: 8). Die Autoren halten aufgrund der begrifflichen Unschärfen Szenarien für möglich, in denen unter dem Label „Diamond" Angebote organisiert werden, die wegen der Nutzung von Drittanbietern wieder in kommerzielle Abhängigkeiten und zu Preissteigerungen führen (ebd.: 6). Sie führen als weitere Argumente gegen DOA an, dass ein Ausweichen auf nicht etablierte Kategorien nicht wünschenswert ist und es unwahrscheinlich sei, dass die Zielgruppe der Wissenschaftlerinnen und Wissenschaftler den Begriff aufnähmen. Außerdem komme in den Open-Access-Policies DOA nicht vor und in der Publikationsberatung könne es zu Kommunikationsproblemen kommen (ebd.: 10f).

Sicher ist es so, dass „Gold Open Access" inzwischen als Begriff etabliert ist, sowohl bei den Forschenden als auch bei Verlagen und Forschungsförderern. Missverständnisse gibt es in der Publikationsberatung dennoch regelmäßig, da Verlage Gold OA auf Artikelebene vermarkten, während die finanzierenden Bibliotheken nach den DFG-Kriterien in der Regel Gold OA nur auf Zeitschriftenebene finanzieren. Mit anderen Worten: unerwünschte Hybrid-Open-Access-Geschäftsmodelle, die nur im Rahmen von Transformationsverträgen finanziert werden, sind Wissenschaftlerinnen und Wissenschaftlern häufig nicht vertraut, besonders, wenn sie als Doktoranden erst Erfahrungen mit dem Open-Access-Publizieren sammeln. Die hochspezialisierte akademische

Diskussion um die Weiterentwicklung des wissenschaftlichen Publikationssystems und Open Access wird Nachwuchswissenschaftlerinnen und -wissenschaftler in der Regel weniger beschäftigen, da sie mit ihrer Forschung und den qualitativen Anforderungen an das Publizieren ausgelastet sind und am Ende oft einfach nur nach Finanzierungsmöglichkeiten suchen oder einfach den Vorgaben ihrer Betreuerinnen und Betreuer folgen.

An DOA wird bemängelt, dass es keine allgemeinverbindlichen Kriterien gibt, anhand derer beurteilt werden könne, ob es sich bei einem Publikationsorgan um DOA handele oder nicht. Als Alternative wird vorgeschlagen, den fünf Prinzipien der Fair Open Access Alliance (FOAA) zu folgen. Um als fair zu gelten, müssen demnach erstens die entsprechende Zeitschrift wissenschaftsgeleitet sein, zweitens die Autorinnen und Autoren die Rechte an ihren Werken behalten, drittens die Publikationen vollständig und unmittelbar Open Access sein, viertens Einreichung und Veröffentlichung nicht unmittelbar mit Kosten verbunden sein sowie fünftens eventuell anfallende Gebühren niedrig, angemessen und transparent sein. Als Alternative zu DOA wird deshalb „Fair Open Access" (FOA) vorgeschlagen (vgl. Dellmann and Voigt 2024). Gebührenfreiheit soll in Abgrenzung zu DOA kein Selbstzweck sein. Nach dem FOA-Prinzip sind auch nicht gewinnorientierte Zeitschriften mit Artikelgebühren denkbar.

Mit dem Konzept „rainbow OA" wird DOA schließlich beiseitegelegt und ein übergeordnetes Prinzip von Open Access entworfen, das gegenüber Geschäftsmodellen indifferent ist („business-model agnostic"). Es geht dabei um ein vollständiges Spektrum von Ansätzen, ein gerechtes und inlusives OA zu erreichen (vgl. Legge 2023).

Nils Taubert dämpft die hohen Erwartungen, die mit DOA verbunden sind und hinterfragt das Ziel der Souveränität: „Open-Access-Modelle bewegen sich immer zwischen den Polen Souveränität und Abhängigkeit, und letztlich bleibt uns nur die Wahl zwischen verschiedenen Arten von Abhängigkeit." (Taubert 2024: 2) Bisher gebe es kein „Proof-of-concept" für große Zeitschriften

(ebd.: 11). Die Untersuchung im Projekt CODRIA habe eine Instabilität gezeigt: unbezahlte Arbeit prekär Beschäftigter, Journale, die eingestellt wurden und Abhängigkeiten, die sichtbar werden. Er verweist auf die Notwendigkeit einer dauerhaften Finanzierung (ebd.: 15f).

Ulrich Herb und Benedikt Schmal halten das DOA-Modell aus dem gleichen Grund für nicht ausreichend verstanden. Es sei unklar, ob DOA billiger sein würde als das gegenwärtige System (vgl. Herb and Schmal 2024). DOA könne zwar für Chancengleichheit bei den Publikationsmöglichkeiten sorgen, nicht jedoch Reputation garantieren.

Heinz Pampel und Dorothea Strecker haben, unabhängig von der Diamond-Diskussion im engeren Sinne, die Rolle der wissenschaftlichen Fachgesellschaften als Akteure im Open Access untersucht. Sie regen an, die Migration der Publikationsorgane der Fachgesellschaften in den Open Access zu fördern und mit nicht-kommerziellen Publikationsdienstleistern zu kooperieren (Pampel and Strecker 2020: 34). Sie sehen auch die DFG-geförderten Fachinformationsdienste (FID) in der Rolle, die Publikationsdienste für die Fachgesellschaften auszubauen. Am Ende stoßen diese Ansätze wieder auf das Finanzierungsproblem, da die FIDs jenseits der befristeten Förderlogik vor der Herausforderung stehen, nachhaltige Lösungen zu finden. Für die Fachgesellschaften stellen sich komplementäre Fragen, um vom kommerziellen auf einen DOA-Pfad zu wechseln. Ein Beispiel: Für Robert Harington, Chief Publishing Editor der American Mathematical Society (AMS), passt DOA durchaus zur Mission der Fachgesellschaft. Er sieht jedoch das Finanzierungsproblem: „this approach requires significant funding" (Harington 2017). Es seien größere Investitionen nötig, als ein Journal neu aufzusetzen. Ein hoher Kapitalstock sei für die längerfristige Finanzierung nötig, um Subskriptionseinnahmen zu ersetzen. Dazu käme die Finanzierung aus den Überschüssen der Verlagstätigkeit. Inzwischen hat die AMS immerhin drei Diamond-Zeitschriften herausgebracht.[12] Die Umstellung des gesamten Zeitschriftenportfolios auf DOA bleibt jedoch offen.

4. Fazit: Konkurrenz zweier Open-Access-Pfade

In der oben dargestellten Diskussion um Alternativen zum kommerziellen Open Access geht es um systemische Fragen, die auch in der Organisationstheorie bewegt werden: Wie kommt man aus Handlungslogiken heraus, die das bisherige System nur fortschreiben, aber nicht verändern? Dobusch und Heimstädt (2021) haben aus organisationstheoretischer Perspektive Open Access als einen Pfad und damit pfadabhängige Entwicklungen und deren Überwindung analysiert. Unter Pfadabhängigkeit verstehen sie eine Situation, in der individuelle Akteure oder Organisationen ihre Fähigkeit verloren haben, zwischen Alternativen zu wählen und folgen damit einer Definition von Jörg Sydow. Die Funktion der Organisationskritik besteht dann u.a. darin, „hidden patterns" zu thematisieren, die gerne verdrängt werden (Jörg Sydow et al. 2009: 702). Im OA-Ökosystem geht es jedoch nicht nur um eine Organisation, sondern um interorganisationale Beziehungen, die in ein komplexes Netz eingebettet sind (Verlage, Fachgesellschaften, Herausgebergremien, Forschungsförderer, Bibliotheken). Diese Beziehungsgeflechte können genauso pfadabhängig sein. Dobusch und Heimstädt verstehen OA in dieser Hinsicht als einen Pfad, der wegen der Kommodifizierung der wissenschaftlichen Publikationen fragil bleibt. Darauf hebt die Kritik an den Transformationsverträgen ab. Sie erhoffen sich eine „Verriegelung des OA-Pfades" durch gemeinwohlorientierte Publikationsinfrastrukturen und z.B. eine Umleitung von Gewinnmargen auf vernachlässigte Literaturgattungen („Diamond Road"). Sie schließen ihre Überlegungen mit dem Gedanken, dass ohne einen Strukturwandel demokratischer Öffentlichkeit DOA eventuell nicht machbar ist. In Bezug auf die Wissenschaftskommunikation braucht es vor allem einen Strukturwandel des Reputationssystems und der Finanzierung – beides hätte der in der Organisationskritik die Funktion von „hidden pattern", an denen gearbeitet werden müsste. Es fehlen bislang große DOA-Journals in den STM-Fächern, die ohne diesen Strukturwandel kaum zu realisieren sind.

Die Vorstellung, dass Erwerbungsbudgets genutzt werden, um gemeinnützige Publikationsinfrastrukturen zu finanzieren, funktioniert oft nicht. Zum einen sind an vielen Einrichtungen die Budgets bereits heute strukturell unterfinanziert. Die DEAL-Verträge verschlingen den größten Teil des Etats oder müssen schon über zusätzliche Mittel der Hochschulleitungen finanziert werden. D.h. für DOA muss parallel zu den Transformationsverträgen, Monografien, Datenbanken etc. Geld bereitgestellt werden. Wissenschaftliche Bibliotheken handeln darüber hinaus in ihrer Erwerbungslogik nach den Grundsätzen öffentlicher Haushalte wie denen der Sparsamkeit und Wirtschaftlichkeit. Das führt dazu, dass auch wünschenswerte Diamond-Finanzierungsmodelle nicht mitgetragen werden, weil die Gegenleistung für die eigene Einrichtung den Finanzabteilungen der Hochschulen oft nicht plausibel darstellbar sind. Bekam man früher für eine Subskription einen Zeitschriftentitel geliefert, sind bei einer konsortialen gemeinnützigen Open-Access-Plattform nicht in gleicher Weise im gleichen Haushaltsjahr auch Veröffentlichungen von Autor:innen der eigenen Einrichtungen nachweisbar. Hat man sich an einer gemeinschaftlich getragenen Publikationsplattform mit einem Mitgliedsbeitrag beteiligt, so kann es passieren, dass die Finanzierung im Folgejahr wieder eingestellt wird, weil keine Veröffentlichungen geplant sind oder nach einem Dreijahreszeitraum der Konsortialvertrag nicht verlängert wird. Das sind konkrete Barrieren, die eine einfache Umschichtung von Erwerbungsmitteln zu Diamond-Etats erschweren. Die Finanzierung von DOA bleibt eine Herausforderung.

Abweichend zur Analyse von Dobusch und Heimstädt sollten zwei Open-Access-Pfade unterschieden werden, weil sie organisatorisch unterschiedlich aufgestellt sind und unterschiedlichen Logiken folgen: (1) Kommerzielles OA, das geprägt ist durch die Gewinnorientierung des Oligopols der Wissenschaftsverlage und den damit verbundenen Sachzwängen und unvollständigem Open Access (Hybrid-Zeitschriften, Buchmarkt). Davon zu unterscheiden wäre (2) das nicht-kommerzielle Diamond Open Access, das von öffentlich finanzierten Infrastrukturen getragen

wird und ebenfalls spezifischen Sachzwängen wie Finanzierungsfragen und Reputationsproblemen geprägt ist, aber vollständiges Gold Open Access ermöglicht. Beide Pfade stehen in Wechselwirkung miteinander und konkurrieren um öffentliche Gelder. Es ist derzeit nicht absehbar, dass bei angespannten Haushalten an den Wissenschaftseinrichtungen der DOA-Pfad parallel zum kommerziellen OA-Pfad in einer Weise nachhaltig ausgebaut wird, so dass er ihn ablösen könnte. Es ist keineswegs ausgemacht, dass die aktuell starke wissenschaftspolitische Unterstützung für DOA in Europa mit Förderprogrammen dazu führt, dass die Abhängigkeit von den Wissenschaftsverlagen, vorrangig in den STM-Fächern, entscheidend sinkt. Das würde die Migration von Editorial Boards großer Zeitschriften des Oligopols erfordern und damit einhergehend die Veränderung des wissenschaftlichen Reputationssystems. Selbst unabhängige Fachgesellschaften wie die AMS tun sich schwer, wie oben gezeigt. Es ist leichter, neue DOA-Zeitschriften zu gründen und dies vor allem in den Geistes- und Sozialwissenschaften sowie in kleinen Fächern. Das kann im Globalen Norden den Diamond-Pfad zwar stärken, muss aber absehbar mit dem kommerziellen Pfad konkurrieren. Der kommerzielle Pfad kann durch seine Anziehungskraft (Mega-Journals, Reputation, Sichtbarkeit der Publikationen) den Diamond-Pfad schwächen, wie am Beispiel Kolumbien zu beobachten ist. Es ist das erste lateinamerikanische Land, das Open-Access-Transformationsverträge mit Elsevier, Springer Nature und Taylor & Francis abgeschlossen hat (vgl. Consorcio Colombia).

Anmerkungen

[1] DFG-Förderprogramm Open Access Publizieren (2011–2022) und das Förderprogramm Open Access Publikationskosten (seit 2022): „Wie unterstützt die DFG Open Access in der Infrastrukturförderung?“, https://www.dfg.de/de/foerderung/foerdermoeglichkeiten/programme/infrastruktur/lis/open-access/infrastrukturfoerderung

[2] Homepage der Plattform SciELO: https://www.scielo.org/en/, eingesehen am 20.09.2025

[3] Vgl. https://old.scielo.org/en/about-scielo/program-publication-model-and-scielo-network/, eingesehen am 20.09.2025
[4] Vgl. https://scielo.org/en/about-scielo/scielo-network/, eingesehen am 20.09.2025
[5] Homepage der Plattform Redalyc: https://redalyc.org, eingesehen am 20.09.2025
[6] Vgl. den spanischsprachigen Wikipedia-Artikel zu AmeliCA: https://es.wikipedia.org/wiki/AmeliCA, eingesehen am 20.09.2025
[7] Vgl. DiPP: https://www.hbz-nrw.de/produkte/open-access/dipp/, eingesehen am 20.10.2025
[8] Vgl. https://www.unesco.org/en/articles/announcing-global-diamond-open-access-alliance, eingesehen am 10.09.2025
[9] Vgl. https://diamas.org/about, eingesehen am 09.09.2025
[10] Vgl. die Projekt-Website https://diamasproject.eu, eingesehen am 09.09.2025
[11] Vgl. https://www.craft-oa.eu, eingesehen am 09.09.2025
[12] Vgl. American Mathematical Society (AMS): Open Access Journals https://www.ams.org/publications/journals/open-access/open-access, eingesehen am 22.11.2025

Literatur

Aguado-López, E.; Becerril-García, A.; Macedo-García, A. (2024) Redalyc and AmeliCA. An open infrastructure for advancing science as a global public good. https://doi.org/10.5281/ZENODO.11123534.

Alexander von Humboldt Foundation, German Academic Exchange Service, Deutsche Forschungsgemeinschaft, Fraunhofer Society, Helmholtz Association of German Research Centres, German Rectors' Conference, Leibniz Association, German National Academy of Sciences Leopoldina, Max Planck Society, German Science and Humanities Council (2025) Strategie der Allianz für die Weiterentwicklung des wissenschaftlichen Publikationswesens 2026–2030. https://doi.org/10.5281/ZENODO.15853225.

Ancion, Z.; Borrell-Damián, L.; Mounier, P.; Rooryck, J.; Saenen, B. (2022) Action Plan for Diamond Open Access. https://doi.org/10.5281/ZENODO.6282403.

Becerril García, A. (2024) Science as a Public Good: The Role of Community and Digital Technologies for Equity, Sustainability and

Development in the Global South. Annals of Library and Information Studies 71. https://doi.org/10.56042/alis.v71i4.14248.

Berger, M. (2021) Bibliodiversity at the Centre: Decolonizing Open Access. Development and Change 52, 383–404. https://doi.org/10.1111/dech.12634.

Björk, B.-C. (2017) Growth of hybrid open access, 2009–2016. PeerJ 5, e3878. https://doi.org/10.7717/peerj.3878.

Bosman, J.; Frantsvåg, J.E.; Kramer, B.; Langlais, P.-C.; Proudman, V. (2021) OA Diamond Journals Study. Part 1: Findings. Zenodo. https://doi.org/10.5281/ZENODO.4558704.

Dellmann, S.; Van Edig, X.; Rücknagel, J.; Schmeja, S. (2022) Facetten eines Missverständnisses. o-bib. Das offene Bibliotheksjournal / Herausgeber VDB Bd. 9 Nr. 3. https://doi.org/10.5282/O-BIB/5849.

Dellmann, S., Voigt, M. (2024) Fair Open Access fördern. Front Matter. https://doi.org/10.64395/v0d11-y2804.

Deutsche Forschungsgemeinschaft, AG Publikationswesen (2022) Wissenschaftliches Publizieren als Grundlage und Gestaltungsfeld der Wissenschaftsbewertung. https://doi.org/10.5281/ZENODO.6538163.

Dobusch, L.; Heimstädt, M. (2021) „Strukturwandel der wissenschaftlichen Öffentlichkeit. Konstitution und Konsequenzen des Open-Access-Pfades", in: Ein Neuer Strukturwandel Der Öffentlichkeit? Nomos Verlagsgesellschaft, S. 425–454. https://doi.org/10.5771/9783748912187-425.

Else, H. (2018) „Radical open-access plan could spell end to journal subscriptions", in: Nature 561, S. 17–18. https://doi.org/10.1038/d41586-018-06178-7.

Farley, A.; Langham-Putrow; A., Shook, E.; Sterman, L.; Wacha, M. (2021) Transformative agreements: Six myths, busted: Lessons learned. College Research Libraries News 82. https://doi.org/10.5860/crln.82.7.298.

Fokusgruppe Scholar-Led.Network (2021) Das scholar-led.network-Manifest. https://doi.org/10.5281/ZENODO.4925784.

Fuchs, C.; Sandoval, M. (2013) The Diamond Model of Open Access Publishing: Why Policy Makers, Scholars, Universities, Libraries, Labour Unions and the Publishing World Need to Take Non-Commercial, Non-Profit Open Access Serious. tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society 11, S. 428–443. https://doi.org/10.31269/triplec.v11i2.502.

Gemeinsame Task Group Open-Access-Transformation der Arbeitskreise Open Science und Bibliotheks- und Informationsmanagement und dem Helmholtz Open Science Office (2024) Memorandum zur Open-Access-Transformation bei der Helmholtz-Gemeinschaft – Positionspapier der Arbeitskreise Open Science und Bibliotheks- und Informationsmanagement. Helmholtz Open Science Office. https://doi.org/10.48440/OS.HELMHOLTZ.078.

Ginsparg, P. (2011) It was twenty years ago today .... https://doi.org/10.48550/ARXIV.1108.2700.

Sydow, J.; Schreyögg, G.; Koch, J. (2009) „Organizational Path Dependence: Opening the Black Box", in: The Academy of Management Review 34, 689–709.

Mittermaier, B. (2025) Transformationsverträge sind eine Sackgasse. https://doi.org/10.5282/O-BIB/6117.

Mittermaier, B. (2015) Double Dipping beim Hybrid Open Access – Chimäre oder Realität? Informationspraxis Bd. 1. https://doi.org/10.11588/IP.2015.1.18274.

Nötzelmann, C. (2007) „Die Initiative Digital Peer Publishing NRW – Erfahrungen mit dem kollaborativen Publikations- und Kommunikationsmodell", in: Open Innovation – Neue Perspektiven im Kontext von Information und Wissen: 10. Internationales Symposium für Informationswissenschaft, Schriften zur Informationswissenschaft. UVK Verlag, S. 469–478.

Pampel, H.; Jahn, N.; Bertelmann, R.; Horstmann, W.; Rothfritz, L.; Ferguson, L.M.; Schmidt, B.; Stisser, A. (2024) Datenpraxis zur Gestaltung der Open-Access-Transformation – Analyse, Empfehlung, Training & Vernetzung (OA Datenpraxis). https://doi.org/10.5281/zenodo.10794298.

Pampel, H.; Strecker, D. (2020) Wissenschaftliche Fachgesellschaften und Open Access in Deutschland. https://doi.org/10.2312/OS.HELMHOLTZ.009.

Šimukovič, E. (2024) Transformative Agreements Are a Blind Alley. Katina Mag. https://doi.org/10.1146/katina-20241008-1.

Smits, R.-J.; Pells, R. (2022) Plan S for Shock: Science. Shock. Solution. Speed. Ubiquity Press. https://doi.org/10.5334/bcq.

Stäcker, T., Apel, J., Arning, U., Burschel, P., Christof, J., Effinger, M., Elsner, C., Finger, J., Günther, A.-C., Hagedoorn, J.M., Jansky, C., Kaiser, M., Meinecke, I., Mischke, D., Pieper, D., Riesenweber, C., Rißler-Pipka, N., Schmitz, J., Schobert, D., Schubotz, M., Siegert, O.,

Strötgen, R., Teschke, O., Tochtermann, K., Wermbter, K., Wiegand, M., Wonke-Stehle, J., Wrzesinski, M., Wuttke, U., Zepf, R., Ziegler, B. (2025) SeDOA – Servicestelle Diamond Open Access. https://doi.org/10.5281/ZENODO.15043760.

Taubert, N. (2024) Diamond OA: Vierte und finale Welle der Open-Access-Förderung? https://doi.org/10.5281/ZENODO.13768808.

Tautz, D.; Holzer, A.; Schmidt, K.; Buchner, J.; Grötschel, M.; Jurburg, S. (2025) Ein neues Verfahren zur direkten Finanzierung wissenschaftlicher Publikationen, Diskussion. MyCoRe Community. https://doi.org/10.26164/LEOPOLDINA_03_01261.

UNESCO (2021) UNESCO Recommendation on Open Science. UNESCO. https://doi.org/10.54677/MNMH8546.

Wissenschaftsrat (2022) Empfehlungen zur Transformation des wissenschaftlichen Publizierens zu Open Access. https://doi.org/10.57674/FYRC-VB61.

Internetquellen

AmeliCA. https://es.wikipedia.org/wiki/AmeliCA, eingesehen am 20.09.2025.

American Mathematical Society (AMS): Open Access Journals. https://www.ams.org/publications/journals/open-access/open-access, eingesehen am 22.11.2025.

Bundesministerium für Bildung und Forschung (BMBF): Open Access in Deutschland. Gemeinsame Leitlinien von Bund und Ländern, Stand: 05.06.2023. https://www.bmftr.bund.de/SharedDocs/Publikationen/DE/1/772960_Open_Access_in_Deutschland.pdf?__blob=publicationFile&v=5, eingesehen am 12.09.2025.

Consorcio Colombia: Science now closer to Colombia and from Colombia to the world: Consortium Colombia reaches first transformative agreements in Latin America. https://www.consorciocolombia.co/science-now-closer-to-colombia-and-from-colombia-to-the-world-consortium-colombia-reaches-first-transformative-agreements-in-latin-america/, eingesehen am 30.11.2025.

CRAFT-OA. https://www.craft-oa.eu, eingesehen am 09.09.2025.

Cressey, D.: Open access: are publishers 'double dipping'? Nature archived newsblog, 20.10.2009. https://blogs.nature.com/news/2009/10/story_landis_resigns_from_auti.html, eingesehen am 02.11.2025.

DIAMAS. https://diamas.org/about, eingesehen am 09.09.2025.
DIAMAS: About. https://diamas.org/about, eingesehen am 09.09.2025.
DIAMAS: Welcome to the DIAMAS Project. https://diamasproject.eu, eingesehen am 09.09.2025.
DiPP. https://www.hbz-nrw.de/produkte/open-access/dipp/, eingesehen am 20.10.2025.
Eve, M. P.: Diamond Mining, 31.03.2021. https://www.coalition-s.org/blog/diamond-mining/, eingesehen am 11.11.2025.
Harington, R.: Diamond Open Access, Societies and Mission. The Scholarly Kitchen, 01.06.2017. https://scholarlykitchen.sspnet.org/2017/06/01/diamond-open-access-societies-mission/, eingesehen am 10.11.2025.
Herb, U., Schmal, B.: The benefits of diamond are not crystal clear. Research Professional News, 26.09.2024. https://www.researchprofessionalnews.com/rr-news-europe-views-of-europe-2024-9-the-benefits-of-diamond-are-not-crystal-clear/.
Legge, M.: Report from Equity in Open Access workshop #1: the APC debate, reflections and rainbows, 21.03.2023. https://www.oaspa.org/news/report-from-equity-in-open-access-workshop-1-the-apc-debate-reflections-and-rainbows/, eingesehen am 02.11.2025.
Manifesto on Science as Global Public Good: Noncommercial Open Access, 1st Global Summit on Diamond Open Access, Toluca, Mexico / October 27, 2023. https://globaldiamantoa.org/manifiesto/#/, eingesehen am 30.11.2025.
National Library of Sweden: The Bibsam Consortium's new action plan, 26.02.2024. https://www.kb.se/for-bibliotekssektorn/eng/news-from-the-national-library/news/2024-02-26-the-bibsam-consortiums-new-action-plan.html, eingesehen am 30.09.2025.
Open Research Europe (ORE). https://open-research-europe.ec.europa.eu, eingesehen am 28.11.2025.
Plan S: cOAlition S confirms the end of its financial support for Open Access publishing under transformative arrangements after 2024, 26.01.2023. https://www.coalition-s.org/coalition-s-confirms-the-end-of-its-financial-support-for-open-access-publishing-under-transformative-arrangements-after-2024/, eingesehen am 15.11.2025.
Prosser, D.: The costs of double dipping. RLUK Blog – Research Libraries UK, 06.02.2015. https://www.rluk.ac.uk/the-costs-of-double-dipping/, eingesehen am 21.11.2025.

Rat der Europäischen Union: Schlussfolgerungen zu Wegen des hochwertigen, transparenten, offenen, vertrauenswürdigen und fairen wissenschaftlichen Publizierens, Brüssel, 23.05.2023. https://data.consilium.europa.eu/doc/document/ST-9616-2023-INIT/de/pdf, eingesehen am 25.09.2025.

Redalyc. https://redalyc.org, eingesehen am 20.09.2025.

Schwamm, H.: Open Voices Interview with Wilhelm Widmark on Open Access in Sweden. The HardiBlog, 11.09.2023. https://hardimanlibrary.blogspot.com/2023/09/wilhelmwidmark.html, eingesehen am 22.11.2025.

SciELO. https://www.scielo.org/en/, eingesehen am 20.09.2025.

SciELO Network. https://scielo.org/en/about-scielo/scielo-network/, eingesehen am 20.09.2025.

SciELO: Program, Publication Model, and SciELO Network. https://old.scielo.org/en/about-scielo/program-publication-model-and-scielo-network/, eingesehen am 20.09.2025.

Toluca – Cape Town Declaration on Diamond Open Access, 2nd Global Summit on Diamond Open Access, 8–14 December 2024. https://doasummit.uct.ac.za/, eingesehen am 07.09.2025.

UNESCO: Announcing the Global Diamond Open Access Alliance. https://www.unesco.org/en/articles/announcing-global-diamond-open-access-alliance, eingesehen am 10.09.2025.

KAPITEL 4

Datentracking – Gefahr für die Wissenschaft?[1]

Bernhard Mittermaier

1. Einleitung

Online-Dienste wie Social-Media-Plattformen und Smartphone-Apps werden schon oft ohne monetäre Bezahlung angeboten. Unternehmen, die hinter diesen „kostenlosen" Angeboten stehen, erwirtschaften dennoch zum Teil erhebliche Umsätze und Gewinne, die u.a. durch die Weiterverwendung von Nutzerdaten generiert werden. So werden z.B. die bei der Nutzung ermittelten persönlichen Interessen für maßgeschneiderte Werbung genutzt und diese Werbeplätze im Verfahren des „Real-Time-Bidding"[2] entsprechend vergütet. Die Nutzenden bezahlen die Angebote mit ihren Daten, die persönlichen Informationen werden somit zur Währung (Schreiber 2021). Seit 2022 ist diese Art der Bezahlung auch ausdrücklich im Bürgerlichen Gesetzbuch (BGB § 312 Abs. 1a) geregelt: die Weitergabe von Daten ist einer Geldzahlung gleichgesetzt. Bereits hier wird das Prinzip der Datensouveränität,

Zitationsvorschlag für dieses Kapitel:
Mittermaier, B. 2026. Datentracking – Gefahr für die Wissenschaft? In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* S. 69–95. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.e. License: CC BY 4.0

also die selbstbestimmte Kontrolle über Erhebung, Speicherung, Nutzung und Verarbeitung der eigenen Daten (Augsberg & Gehring 2022), zumindest tangiert, denn eine Teilhabe am politischen und gesellschaftlichen Leben ohne die Nutzung solcher Dienste ist für sehr viele sicher schwer vorstellbar, die Preisgabe der Daten mithin nicht mehr völlig freiwillig.[3]

Wenn sich Wissenschaftsverlage die Nutzung ihrer Online-Plattformen beispielsweise zur wissenschaftlichen Literaturrecherche mit Nutzerdaten bezahlen lassen, dann ist die Freiwilligkeit der Datenpreisgabe sicher nicht mehr gegeben. Es ist ein konstitutiver Bestandteil der Arbeit von Wissenschaftler:innen, wissenschaftliche Publikationen zu lesen, und wenn dies in elektronischer Form erfolgt, so geht es mit einer zwangsweisen Preisgabe ihrer Daten einher. Die Datensouveränität von Wissenschaftler:innen ist somit klar verletzt. Oftmals geht dieses „Datentracking in der Wissenschaft" völlig geräuschlos und ohne ausreichende Information den Nutzenden vonstatten. Ein zweiter Aspekt: Bei den oben genannten Angeboten ist es ja gerade das Geschäftsmodell, zu Gunsten der Datenpreisgabe auf eine monetäre Vergütung zu verzichten. Die Wissenschaftsverlage werden dagegen entweder für das Lesen (Subskriptionsverträge) oder für das Schreiben (Publikationsgebühren) oder für beides in einem Vertrag (Transformationsverträge) bezahlt. Sie lassen sich also mit Nutzerdaten ein zweites Mal bezahlen.

Dieser Beitrag diskutiert zunächst, warum Verlage mutmaßlich Daten erheben und wie sie dabei vorgehen, erörtert anschließend juristische Aspekte und die Reaktionen aus der Wissenschaft auf dieses Vorgehen. Im zweiten Teil wird die aktuelle Praxis und Handhabung des Datentracking-Themas durch Wissenschaftsverlage erörtert, insbesondere in den 2023 abgeschlossenen DEAL-Verträgen.

2. Motive des Datentracking

Beim Datentracking in der Wissenschaft (Wissenschaftstracking) werden „die Daten von Wissenschaftlerinnen und Wissenschaftlern

[...] bei der Nutzung von Informationsdiensten wie z.B. der Literaturrecherche getrackt, das heißt festgehalten und gespeichert." (DFG-Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme 2021). Dies kann im unmittelbaren Interesse der Nutzenden liegen (z.B. Speicherung und Nutzung einer Email-Adresse zur Bearbeitung einer Anfrage), in einem zumindest mittelbaren Interesse der Nutzenden (z.B. Erstellung von Nutzungsstatistiken[4] für die Einrichtung der Nutzenden) oder in einem unbestrittenen Interesse der Verlage (z.B. Speicherung des Bereichs der IP-Adressen einer lizenzierenden Einrichtung bei zugangsbeschränkten Angeboten). In der Kritik sind nicht die zuvor genannten Zwecke der Datenspeicherung, sondern der Umstand, dass Verlage oftmals „alles" speichern, und zwar über längere Zeiträume hinweg, ggf. Daten aus verschiedenen Nutzungszusammenhängen kombinieren und so Personenprofile anlegen. Diese Sammlung, Aggregation und/oder Weitergabe und Nutzung der aufgezeichneten Daten(-spuren) und persönlichen Informationen geschieht oft auf intransparente und unregulierte Weise. Warum tun dies Verlage? Der Verlag Frontiers schreibt dazu in seiner Privacy Policy: „We combine information you have provided us and information coming from different sources [...] to have a better understanding of your publishing activities, academic and scientific background and areas of expertise and of your use of the Services. These sources of information include your activity as user of the Services, as well as publicly available sources, including your publications on public-facing social media platforms. We create a profile of your publishing activities, including citations, impact, affiliation and awards. We use this information to assess manuscripts submitted for publication in Journals, or to select and bring to your attention the right cooperation opportunities which are relevant to you and match your interests and expertise." (Frontiers Media SA 2023)

Die Möglichkeit zur direkten Monetarisierung der Nutzungsdaten hat beispielsweise der Verlag Elsevier im Kontext seines Produkts „Scopus AI".[5] Diese gegen Aufpreis angebotene Erweiterung der Datenbank Scopus um Methoden der künstlichen

Intelligenz ermöglicht es u.a., Namensvorschläge für Expert:innen zu einem bestimmten Thema zu erhalten. Ganz anders gelagert wäre ein anderer „Grund" für Datentracking: Wie die meisten Onlineanbieter verwenden auch Verlage zur Erstellung ihrer Webseiten Baukästen, die Trackingtools und Einstellungen oftmals per default enthalten. Es wäre also auch denkbar, dass Verlage Datentracking unwissentlich betreiben und die gesammelten Daten überhaupt nicht verwenden. Ohne die juristische Frage zu diskutieren, wer in einem solchen Fall die Verantwortung trägt, ist aber mindestens festzuhalten, dass es nach den DEAL-Verhandlungen 2022/2023 eine unschuldige Unkenntnis des Datentracking auf Seiten der Wissenschaftsverlage nicht mehr geben kann.

3. Techniken des Datentracking

Nachstehend wird ein Überblick zu verschiedenen Techniken des Datentracking gegeben. Für eine vertiefte Darstellung sei auf die Publikation (Altschaffel et al. 2024) und die dort zitierte weiterführende Literatur verwiesen.

Die Erfassung von Daten erfolgt üblicherweise während der Verwendung eines Browsers, aber auch während der Nutzung von Apps oder E-Mail. Dabei gibt es sowohl direkte Datenflüsse an den Verlag (First-Party-Tracking) als auch Datenflüsse an Dritte, z.B. an Buchungs- oder Zahlungssysteme, an Cloud-Anbieter wie Amazon oder Cloudflare, oder an Werbesysteme z.B. von Google oder Adobe (Third-Party-Tracking). Die Einbindung solcher Dritter ist für den Nutzenden nicht immer transparent. So ist beispielsweise die Einbindung von Drittdiensten zur Bereitstellung bestimmter Schriftarten („Google Fonts") üblich und führt dazu, dass der Drittdienst zwangsläufig aufgerufen wird und dadurch Daten übermittelt werden. Oft ist es nur eingeschränkt oder gar nicht möglich, diese Art des Tracking zu verhindern (Bettinger et al. 2023).

Tracking kann in verschiedenen Ausprägungen stattfinden, z.B.:

- *Tracking von Zugriffen auf Webseiten und deren Inhalten mit personalisierten Parametern wie Cookies.* Dabei wird z.B. übermittelt, wer wann ein bestimmtes Dokument einsieht.
- *Tracking innerhalb von Dokumenten z.B. mit zusätzlichen Softwarebausteinen wie Zählpixeln.* Dazu werden regelmäßig Daten übermittelt, damit die Gegenseite fortlaufend darüber informiert ist, ob ein Dokument noch und von wem eingesehen wird (Kiltz et al. 2023).
- *Identifikation durch Analyse der Systemkonfiguration:* Die auf einem Rechner vorhandene Kombination von Betriebssystemversion, Sprache, installierten Programmen etc. ist in der Summe meist einzigartig und somit genau einem Rechner zuordenbar. Wenn man sich von diesem Rechner aus nur einmal mit Klarnamen an einem System anmeldet, kann diese Verbindung fortan stets hergestellt werden. Beispielsweise setzt Frontiers dieses „device fingerprinting" ein.[6]
- *Personalisierung der genutzten Dokumente:* Eine pdf-Datei wird mit einem sichtbaren oder unsichtbaren Wasserzeichen versehen, das z.B. Tag und Uhrzeit des Abrufs sowie die verwendete IP-Adresse enthält und somit eine Nachverfolgung ermöglicht, wenn eine Publikation, die eigentlich hinter einer Paywall liegt, frei im Internet verfügbar ist. Beispielsweise erfolgt dies beim Verlag Nomos auch bei Open-Access-Artikeln.

Dass all dies (und noch mehr) auch faktisch geschieht, zeigt z.B. die Datenschutzrichtlinie von Wiley: „We collect this information through tracking technologies, including first and third-party cookies, web beacons (also known as tracking pixels or pixel tags), embedded scripts, location-identifying technologies, device recognition technologies, and session replay software. We may combine this information with other information that we hold." (John Wiley & Sons Inc. 2024)[7] Die genannten Technologien werden außerdem ständig weiterentwickelt und ergänzt, z.B. um Scripting-Technologien[8] auf den jeweiligen Webseiten (Lindstrot 2024).

Was die Beobachtung der Datenübermittlung von außen keinesfalls leisten kann, ist dem Getrackten Aufschluss darüber zu geben, ob und wie lange diese Daten gespeichert, in welchen Ländern sie verarbeitet, ob sie an Dritte weitergeleitet werden usw. Ganz allgemein: Es kann am ehesten festgestellt werden, *dass* Daten erhoben werden; wesentlich schwieriger ist es festzustellen,

welche Daten erhoben werden. Was im Anschluss mit den erhobenen Daten geschieht, für welche Analysen und Produktentwicklungen sie weiterverwendet werden, ist nicht festzustellen. Da die Verlage in ihren Datenschutzrichtlinien ebenso wie im direkten Gespräch allenfalls sehr generische Angaben zum *Zweck* und zur *Art* der Verwendung machen, kann aus dieser Perspektive kein auch nur halbwegs erschöpfender Befund erhoben werden. Eine echte Handhabe besitzen nur die einzelnen Nutzenden selbst durch Wahrnehmung ihres Rechts auf Auskunft ihrer gesammelten Daten (Art. 15 DSGVO) sowie die Datenschutzbehörden im Rahmen ihrer gesetzlichen Aufgaben und Befugnisse.

4. Rechtlicher Rahmen für Datentracking

Die Datenschutzgrundverordnung (DSGVO) setzt der Erhebung, Speicherung und Verarbeitung von Daten und somit auch dem Datentracking einen zwingend gültigen rechtlichen Rahmen. So gilt z.B. gemäß Art. 5 Abs. 1 DSGVO:

> „Personenbezogene Daten müssen
> (a) auf rechtmäßige Weise, nach Treu und Glauben und in einer für die betroffene Person nachvollziehbaren Weise verarbeitet werden („Rechtmäßigkeit, Verarbeitung nach Treu und Glauben, Transparenz“);
> (b) für festgelegte, eindeutige und legitime Zwecke erhoben werden und dürfen nicht in einer mit diesen Zwecken nicht zu vereinbarenden Weise weiterverarbeitet werden [...] („Zweckbindung“);
> (c) dem Zweck angemessen und erheblich sowie auf das für die Zwecke der Verarbeitung notwendige Maß beschränkt sein („Datenminimierung“); [...]
> (d) in einer Form gespeichert werden, die die Identifizierung der betroffenen Personen nur so lange ermöglicht, wie es für die Zwecke, für die sie verarbeitet werden, erforderlich ist [...] („Speicherbegrenzung“);[...].“[9]

Wie weiter unten in den Ausführungen zu den DEAL-Verträgen gezeigt wird, bestehen bei vielen Verlagen erhebliche Zweifel

daran, dass diese Anforderungen erfüllt sind. So kann der Autor selbst nach monatelanger Beschäftigung mit der Thematik für sich nicht nachvollziehen, wie seine Daten verarbeitet werden (lit a.).

Zu den weiteren zwingenden Anforderungen zählt, dass alle Betroffenen stets umfangreich zu informieren sind (Art. 12–14 DSGVO) und dass hohe Anforderungen an die Sicherheit der Datenverarbeitung greifen (Art. 33 DSGVO). Zentral ist, dass die Verarbeitung personenbezogener Daten ausdrücklich erlaubt sein muss. Die aus Verlagssicht niedrigste Hürde ist das Vorliegen eines „berechtigten Interesses" des Verlags oder eines Dritten (Art. 6 Abs. 1 lit. f DSGVO). „Berechtigtes Interesse" ist ein unbestimmter Rechtsbegriff, zu dem die Erwägungsgründe 47–49 der DSGVO[10] immerhin Erläuterungen geben. Die „Artikel 29 Arbeitsgruppe" der EU führt in einer Stellungnahme zu Richtlinie 95/46/EG[11] aus: „In Artikel 7 Buchstabe f ist die letzte von sechs Voraussetzungen für die gesetzlich zulässige Verarbeitung personenbezogener Daten aufgeführt. Im Klartext wird darin eine Prüfung der Ausgewogenheit zwischen dem berechtigten Interesse des für die Verarbeitung Verantwortlichen oder Dritter, denen die Daten übermittelt werden, und den Interessen oder Grundrechten der betroffenen Person gefordert. Das Ergebnis dieser Prüfung der Ausgewogenheit der Interessen entscheidet darüber, ob Artikel 7 Buchstabe f als Rechtsgrundlage für eine Verarbeitung geltend gemacht werden kann." (Artikel-29-Datenschutzgruppe 2014) Es ist unklar, inwiefern diese Prüfung zwischen den Rechten betroffener Personen, d.h. den Nutzenden, und den „berechtigten Interessen" des Verlages faktisch stattfindet.

In anderen Fällen ist Datentracking nur zulässig, wenn die betroffene Person zuvor ihre ausdrückliche und freiwillige Einwilligung gegeben hat (Art. 6 Abs. 1 lit. a DSGVO). An der Ausdrücklichkeit fehlt es unter anderem, wenn die Einwilligung durch Ankreuzen von Kästchen erteilt wird (Erwägungsgrund 32 DSGVO). An der Freiwilligkeit fehlt es bereits, wenn die Bereitstellung eines Dienstes an eine obligatorische Einwilligung geknüpft ist (Art. 7 Abs. 4 DSGVO). Schließlich gelten noch strengere Anforderungen,

wenn die Daten in ein Land außerhalb der EU übermittelt werden sollen (Art. 45 ff. DSGVO).

Relevant sind diese Themen aber nicht nur für die von der Datenverarbeitung betroffenen Personen, sondern möglicherweise auch für die Bibliotheken und wissenschaftlichen Einrichtungen, die Vertragspartner der Verlage sind. Es ist anzunehmen, dass eine gemeinsame Verantwortung im Sinne von Art. 26 DSGVO vorliegt (Beurskens 2024). Aus Urteilen des Europäischen Gerichtshofs[12] kann man ableiten, dass diese gemeinsame Verantwortung bereits dann vorliegt, wenn die Bibliothek/Wissenschaftseinrichtung (1) die Datenverarbeitung durch die Verlage ermöglicht und (2) selbst von der Datenverarbeitung (auch nur mittelbar) profitiert. Bibliotheken vermitteln für ihre Einrichtungen den Zugang der Forschenden zu Verlagsplattformen, wodurch erst deren personenbezogene Daten durch die Verlage erhoben, gespeichert und genutzt werden können. Die Bibliotheken erhalten von den Verlagen Statistiken, die teilweise nur auf Grundlage einer Identifikation der Nutzenden erhoben werden können, also das Datentracking voraussetzen. Bei der Verifikation der Autor:innen[13] arbeiten Verlage und Bibliotheken eng zusammen, oft in einem gemeinsam genutzten Dashboard. Wenn eine solche gemeinsame Verantwortung vorliegt, können gemäß Art. 26 Abs. 3 DSGVO die betroffenen Personen Ansprüche gegen jeden Verantwortlichen geltend machen. Die Bibliothek würde in diesem Fall – jedenfalls zunächst – gegenüber betroffenen Personen aus ihrer Einrichtung auch für Verstöße des Verlags haften.

5. Reaktionen von Wissenschaft und Bibliotheken

Seit einigen Jahren gibt es eine Gegenbewegung aus Wissenschaft und Bibliotheken zu den Data Analytics-Ansätzen der Verlage und den damit verbundenen Datentracking-Aktivitäten (Schonfeld 2018). Sie dient letztlich dem Schutz des Grundrechts auf Wissenschaftsfreiheit und des Rechts auf informationelle Selbstbestimmung. International ist beispielsweise die frühe Publikation von

Cody Hanson zu nennen (Hanson 2019), die auf Drittanbieter-Code auf Verlagswebseiten hinwies. Es folgten Veröffentlichungen von SPARC zu Überwachungsmechanismen der Verlage im Allgemeinen (SPARC 2021) und von Elsevier (Yoose & Shockey 2023) und Springer Nature (Yoose & Shockey 2024) im Speziellen, die Charakterisierung von Verlagen als „Überwachungsunternehmen" (Pooley 2022) und die Monographie „Data Cartels" (Lamdan 2022), in der insbesondere die Aktivitäten von RELX (Mutterkonzern von Elsevier) und Thomson Reuters (u.a. Anbieter von Datenbanken) untersucht wurden. Die US-amerikanische Menschenrechtsorganisation American Civil Liberties Union hat sich klar gegen derartige Aktivitäten positioniert.

In Deutschland ist u.a. das Engagement von Renke Siems (Siems 2021; 2022) und Gerhard Lauer (Lauer 2022) hervorzuheben, die sowohl wissenschaftlich als auch für ein allgemeineres Publikum über Datentracking in der Wissenschaft publiziert haben. U.a. Björn Brembs und Konrad Förstner haben 2021 die internationale Initiative „Stop Tracking Science"[14] ins Leben gerufen. Nach einer Beschäftigung des Ausschusses für Wissenschaftliche Bibliotheken und Informationssysteme (AWBI) der DFG mit dem wissenschaftspolitisch relevanten Themenkomplex hat dieser 2021 ein Informationspapier veröffentlicht (DFG-Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme 2021), welches die gegenwärtigen Praktiken des Datentracking darstellt, mögliche Konsequenzen für die Wissenschaft benennt und zur Klärung der rechtlichen, technischen und ethischen Rahmenbedingungen anregt. Es folgte eine Aufarbeitung des Themas in der Zeitschrift ABI Technik mittels drei Beiträgen (Freiberg 2022; Reuter & Söllner 2022; Roßnagel 2022) und einer Replik von Springer Nature (Springer Nature Group 2022). Der AWBI hat im Weiteren eine Arbeitsgruppe eingerichtet, die das Datentracking-Thema aufgreifen und in die DEAL-Verhandlungen einbringen sollte. Ziel der Arbeit war, zur Transparenz der Datentracking-Aktivitäten der Verlage beizutragen, Datenschutzaspekte in den kommenden DEAL-Verträgen zu verankern und dadurch eine Mitbestimmung der Wissenschaften zu gewährleisten. Bei der größten deutschen

Bibliothekskonferenz BiblioCon wurde das Thema 2023 in einer Podiumsdiskussion behandelt (Bilic-Merdes & Lauer 2023), die DEAL-Verhandlungen wurden mittels Rechtsgutachten und durch die kontinuierliche Beratung und Begleitung der Arbeitsgruppe unterstützt, und schließlich mündete die Arbeit in einer gemeinsamen Publikation der AWBI/DEAL-Arbeitsgruppe (Altschaffel et al. 2024).

6. Der Praxisfall „DEAL"

6.1 Das Projekt DEAL

Die Zeitschriftenkrise (Meier 2002) in den 1990er Jahren war gekennzeichnet durch stark steigende Preise insbesondere bei den Zeitschriften im STM-Bereich (Science, Technology and Medicine). Annähernd gleichbleibende Bibliotheksetats führten zu Abbestellungen von wissenschaftlichen Zeitschriften und auch zu Einschränkungen beim Kauf von Monographien und Zeitschriften in den Geistes- und Sozialwissenschaften. Eine Gegenreaktion war die Bildung von Konsortien zur Lizenzierung elektronischer Zeitschriften (Fladung & Dugall 2003), in denen eine Einrichtung (Konsortialstelle) die Verhandlung für viele Bibliotheken in einer Region oder Wissenschaftsorganisation übernimmt. Solche Konsortien entstanden in den Folgejahren in größerer Zahl (Mittermaier & Stanek 2025), auf nationaler Ebene jedoch nicht mit den drei größten Verlagen Elsevier, Springer Nature und Wiley, die zusammen die Hälfte des wissenschaftlichen Publikationsmarktes abdecken. Im August 2013 trat Beate Schücking, damals Rektorin der Universität Leipzig, mit dem Vorschlag an die Öffentlichkeit, Nationallizenzen auch für diese großen Wissenschaftsverlage abzuschließen (Burchard 2013). Auf Vorschlag der Hochschulrektorenkonferenz wurde daraufhin von der Allianz der Wissenschaftsorganisationen zunächst eine Machbarkeitsstudie in Auftrag gegeben und auf deren Basis das Projekt „DEAL" initiiert. Dessen Ziel war es mit den genannten Verlagen Verträge abzuschließen, die den teilnehmenden Einrichtungen sowohl Zugang

zu den Zeitschriften als auch die Möglichkeit zur Open-Access-Publikation ermöglichen sollten. (Mittermaier 2023a) Diese sogenannten „Transformationsverträge" sollen auch dem Ziel der Transformation des wissenschaftliche Publikationswesens in den Open Access dienen (OA2020 2016).

Im Jahr 2016 wurden die DEAL-Verhandlungen mit Elsevier aufgenommen, Springer Nature und Wiley folgten 2017. Die Verhandlungen gestalteten sich sehr schwierig und führten zwischenzeitlich zu einem auch international viel beachteten vertragslosen Zustand mit Elsevier (Vogel 2017). Ab Mitte 2018 waren nahezu alle deutschen Einrichtungen ohne Zugang zu Elsevier-Zeitschriften – ohne nachhaltig negative Auswirkungen für die Wissenschaft (Mittermaier 2023b). Mit Wiley (ab 2019) und Springer Nature (ab 2020) konnten Transformationsverträge geschlossen werden ebenso mit Elsevier im Herbst 2023. Im Zuge der Elsevier-Verhandlungen sowie der Verhandlungen für Anschlussverträge mit Wiley und Springer Nature (ab 2024) wurde das Thema „Datentracking" prominent platziert.

6.2 Datentracking und DEAL

Nachfolgend wird dargelegt, wie in den drei im Jahr 2023 von der MPDL Services gGmbH (MPDLS)[15] abgeschlossenen DEAL-Verträgen – mit Elsevier (Elsevier B.V. & MPDL Services gGmbH 2023), Wiley (Wiley-VCH GmbH & MPDL Services gGmbH 2023) und Springer Nature (Springer Nature Customer Service Center GmbH & MPDL Services gGmbH) – einzelne Datenschutzaspekte geregelt wurden, zunächst in einer Übersicht (Tabelle 4.1).

In allen DEAL-Verhandlungen war die Frage strittig, ob eine *gemeinsame Verantwortlichkeit* der Datenverarbeitung (Art. 26 Abs. 1 S. 1 DSGVO) vorliegt: Während die Jurist:innen der DFG-Arbeitsgruppe und der beteiligten Anwaltskanzlei deren Vorliegen bejahten, wurde dieses von den Jurist:innen aller drei Verlage verneint. Nach Meinung der Verlage sind die Verarbeitungsschritte klar voneinander abgrenzt und eindeutig jeweils einer

Tabelle 4.1: Regelungen ausgewählter Aspekte mit Datenschutzrelevanz in den DEAL-Verträgen.

	Wiley	Elsevier	Springer Nature
Gemeinsame Verarbeitung	Von Verlagen abgelehnt, stattdessen ein Workaround		
Tracking	Erlaubt bei legitimen Interessen		Nur zur Vertragserfüllung
Wiederherstellung von Cookies	Verlag unternimmt angemessene Anstrengungen	Der Verlag unterlässt es	
Speicherdauer Cookies	Keine Begrenzung		Maximal 1 Monat
Datentransfer	Erlaubt gemäß Art. 44 ff DSGVO		Außerhalb EU/Art. 45-Staaten nur mit Zustimmung der MPDLS
Speicherdauer IP-Adresse	Nicht länger als notwendig;		Maximal 48 Stunden

Sphäre (d.h. entweder dem Verlag oder der Forschungseinrichtung) zugeordnet. Der Dissens konnte auch in mehreren Verhandlungsrunden speziell zu dieser Fragestellung nicht ausgeräumt werden. Letztlich einigte man sich auf einen Workaround: In den DEAL-Verträgen wird postuliert, dass keine gemeinsame Verantwortung der Datenverarbeitung vorliegt. Sollte ein Gericht oder eine Datenschutzbehörde ein Vorliegen bejahen, so verpflichten sich die Vertragsparteien, eine Regelung der gemeinsamen Verantwortung zu vereinbaren. Der jeweilige Verlag verpflichtet sich darüber hinaus, die MPDLS und alle Teilnehmereinrichtungen von eventuellen Strafzahlungen freizustellen.

Außer Frage steht, dass zur Umsetzung der Verträge Daten von Nutzenden erfasst, verarbeitet und verwendet werden müssen. Für das, was über das zwingend Notwendige hinaus geschieht – und darunter fällt auch das *Datentracking* – wurde in den DEAL-Verträgen erstmals in nationalen Verlagsverträgen Regelungen getroffen und so zumindest die Transparenz über das, was getrackt

wird, verbessert. Unterbunden wird das Tracking aber nicht generell. Wortgleich zur ohnehin bestehenden gesetzlichen Lage ist Datentracking vertraglich erlaubt, wenn Nutzende selbst zustimmen oder wenn es zur Erfüllung einer vertraglichen oder rechtlichen Verpflichtung notwendig ist. Im Unterschied zu Springer Nature ist bei Elsevier und Wiley das Datentracking darüber hinaus erlaubt, wenn „berechtigtes Interesse" gemäß Art. 6 Abs. 1 (f) DSGVO vorliegt.[16] Es wird zu beobachten und ggf. auch zu überprüfen sein, wie Wiley und Elsevier diese Möglichkeiten nutzen und die juristischen Abwägungen vornehmen. Bei Springer Nature haben Nutzende es selbst in der Hand, die nicht-essentiellen Cookies zu verweigern und dadurch im Ergebnis Datentracking weitgehend zu unterbinden.

Cookies können bekanntlich von Nutzenden selbst per Browsereinstellung automatisiert gelöscht werden. Elsevier und Springer Nature haben sich verpflichtet, keine Techniken zur *Wiederherstellung gelöschter Cookies* einzusetzen. Wiley dagegen hat sich diesbezüglich nur zu „angemessenen Anstrengungen" verpflichtet. Alle Verlage haben sich verpflichtet, eine *Cookie-Management-Plattform* einzurichten, auf der Nutzer:innen Informationen für ihre Entscheidung über Umfang und Art des Datentracking, dem sie zustimmen, finden können. Dabei ist allerdings zu beachten, dass man die Speicherung der sogenannten „essentiellen Cookies", die (angeblich) technisch notwendig sind, nicht ablehnen kann. Eine Wahlmöglichkeit besteht nur hinsichtlich der Zulassung weiterer Cookies z.B. zu Analyse- und Marketingzwecken.

Hinsichtlich der *Speicherdauer von Cookies* sollte man prima vista annehmen, dass eine Speicherung während der Browser-Session (d.h. bis zum Schließen des Internetbrowsers) ausreichend sein sollte („Session-Cookie"). Der Vertrag mit Springer Nature sieht eine Begrenzung auf einen Monat oder kürzer vor; eine derartig klare Verpflichtung fehlt bei Wiley und Elsevier.

Die konkrete Situation hinsichtlich der essentiellen Cookies der Plattformen der DEAL-Verlage und ergänzend dazu von elf weiteren Verlagen, darunter sechs mit besonderer Bedeutung für die Philosophie, war bei einer Stichprobe wie folgt (Tabelle 4.2):

Bei den DEAL-Verlagen sowie bei Brill, MDPI und Nomos konnten die Angaben zur Speicherdauer der für die Funktion

Tabelle 4.2: Charakteristik der essentiellen Cookies bei verschiedenen Verlagen am 15.06.2024. IP-Adresse 134.94.59.79 (IP-Bereich des Forschungszentrums Jülich). Verwendeter Browser: Firefox V115.12.0esr. Die Angaben zu den ersten sechs Verlagen (Brill bis Wiley) wurden der jeweiligen Cookie-Management-Plattform entnommen, die Angaben zu den unteren acht Verlagen (ACS bis Thieme) durch Vergleich der gespeicherten Cookies bei Nutzung der Plattform und nach Schließen und Neustart des Browsers. Es sei ausdrücklich betont, dass hier nur die essentiellen Cookies dargestellt sind, die unvermeidbar sind. Mit der Wahl „alle Cookies akzeptieren" werden wesentlich mehr Cookies auf dem Rechner installiert.

Verlag	**Essential Cookies gesamt**	**Session Cookies**	**Speicherdauer**			
			bis 48 Stunden	**3 bis 300 Tage**	**über 300 Tage**	**Keine Angabe**
Brill	15	7	1	2	5	
Elsevier	27	12		5	10	
MDPI	25	12		3	10	
Nomos	4	2			2	
Springer Nature	22					22
Wiley	7				2	5
ACS	13	7				6
De Gruyter	5	1				4
Franz Steiner Verlag	4	2				2
Frontiers	1					1
Hogrefe	4					4
Klostermann	1	1				
Philosophy Documentation Center	3					1
Thieme	5	3				2

der Verlagsseite essentiellen Cookies der jeweiligen Cookie-Management-Plattform entnommen werden. Bei Wiley fehlte allerdings bei einigen und bei Springer Nature sogar bei allen Cookies die Angabe der Speicherdauer. Folglich kann nicht überprüft werden, ob Springer Nature die vertragliche Beschränkung auf eine Speicherdauer von maximal 48 Stunden tatsächlich einhält. Die übrigen genannten Verlage haben „essentiell notwendige" Cookies mit Speicherdauern von über 300 Tagen (zumeist rund 1 Jahr, bei Brill und MDPI auch bis zu 1.800 Tage und bei Nomos ein Cookie mit unbegrenzter (!) Speicherdauer). Es ist schwer nachzuvollziehen, inwiefern diese Speicherdauer zur Erbringung des angefragten Dienstes notwendig sein soll. Dies gilt erst recht, wenn, wie bei MDPI, zwangsweise Cookies der Social Media-Plattform LinkedIn gesetzt werden. Im Übrigen steht eine über die Session hinausgehende Speicherdauer auch in explizitem Widerspruch zu Betreiberangaben.[17]

Bei den anderen untersuchten Verlagen (in Tabelle 4.2 die unteren acht Verlage, ACS bis Thieme) existieren zwar allgemeine Aussagen über Cookies, aber keine Cookie-Management-Plattform. Ohne sie dürften gesetzlich nur technisch notwendige (essentielle) Cookies gespeichert werden. Um zumindest eine Abschätzung der Aktivitäten vorzunehmen, wurden die Verlagsplattformen im Selbstversuch zunächst genutzt und anschließend wurde in den Firefox-Einstellungen überprüft, wie viele Cookies gespeichert wurden (Spalte „Essential Cookies insgesamt"). Danach wurde der Browser geschlossen und neu gestartet. Die nun noch verzeichneten Cookies zählen zu „unbekannte Speicherdauer" und die Differenz aus beiden Größen zu „Session Cookies".

Insgesamt gibt es aktuell keinen einzigen Verlag, der die Speicherdauer von Essential Cookies durchweg auf die Dauer der Session beschränkt. Dies bedeutet, dass man stets eine unnötig lang sichtbare Datenspur hinterlässt, wenn Nutzende nicht direkt nach ihrem Plattform-Besuch die Browserfunktion zur automatischen Löschung von Cookies beim Schließen des Browsers verwenden.

Der *Datentransfer* in Länder außerhalb der EU, sogenannte Drittstaaten, kann für international strukturierte Verlage von großer Bedeutung sein. Für die Nutzer:innen verschlechtert sich aber

dadurch tendenziell der Datenschutz, weil das Schutzniveau in Drittstaaten niedriger sein kann. Auch ist zu bedenken, dass in manchen Drittstaaten zusätzliche Risiken hinsichtlich Spionage, Diskriminierung oder Missbrauch der Daten bestehen.

Tabelle 4.3 enthält eine Auswertung der Datenschutzerklärungen der gleichen Verlage wie in Tabelle 4.2. Zwei Verlage, der Franz Steiner Verlag (Franz Steiner Verlag 2024) und das Philosophy Documentation Center (Philosophy Documentation Center o.J.) verarbeiten Daten nur am jeweiligen Sitzland Deutschland bzw. den USA, was in beiden Fällen auf Nachfrage per Email bestätigt wurde.

Tabelle 4.3: Angaben zum Datentransfer in Drittstaaten bei verschiedenen Verlagen am 24.07.2024. Die Anordnung erfolgt nach abnehmendem Schutzniveau und ansonsten alphabetisch. Quelle: jeweilige „Datenschutz"-Webseite.

Verlag	**Sitzland**	**Datentransfer**
Franz Steiner Verlag	D	Kein Datentransfer
Philosophy Documentation Center	USA	USA
Hogrefe	D	EU/Art. 45, ansonsten kein Transfer
Klostermann	D	EU/Art. 45, ansonsten kein Transfer
Nomos	D	EU/Art. 45, ansonsten kein Transfer
Springer Nature	D/CH[18]	EU/Art. 45, darüber hinaus: Vorbehalt im DEAL-Vertrag.
ACS	USA	EU/Art. 45, ansonsten keine Angaben
De Gruyter	D	EU/Art. 45, ansonsten keine Angaben
Thieme	D	EU/Art. 45, ansonsten keine Angaben
Brill	NL	Keine Beschränkung
MDPI	CH	Keine Beschränkung
Elsevier	NL	Keine Beschränkung, Länder explizit genannt
Frontiers	CH	Keine Beschränkung, Länder explizit genannt
Wiley	D/USA[19]	Keine Beschränkung, Länder explizit genannt

Es folgt eine Gruppe von Verlagen, bei denen die Verarbeitung mutmaßlich nur in der EU oder in einem Land mit einem Angemessenheitsbeschluss gemäß Art. 45 DSGVO durchgeführt wird (aktuell Andorra, Argentinien, Kanada, Japan, Neuseeland, Südkorea, Schweiz, USA und Vereinigtes Königreich).[20] Hogrefe, Klostermann und Nomos (Hogrefe Verlag GmbH & Co. KG 2024; Nomos Verlagsgesellschaft mbH & Co. KG 2022; Vittorio Klostermann GmbH 2024) haben dies per Email ausdrücklich bestätigt. Springer Nature sieht in der Privacy-Policy (Springer Nature Switzerland AG 2024) zwar Datentransfers auch in weitere Länder grundsätzlich vor, allerdings bietet der DEAL-Vertrag ein faktisches Mitspracherecht, denn ein Datentransfer in ein Drittland bedarf der vorherigen Zustimmung der MPDLS. Bis Juli 2024 hat Springer Nature keine entsprechende Anfrage gestellt. Bei den übrigen Verlagen in dieser Gruppe (ACS, De Gruyter und Thieme) kann man den Datenschutzerklärungen entnehmen, dass Datentransfers in andere EU-Staaten und/oder Länder mit Angemessenheitsbeschluss nach Art. 45 (insbesondere USA) stattfinden (American Chemical Society 2024; Georg Thieme Verlag KG 2021; Walter de Gruyter GmbH 2024). Zu weiteren Transfers gibt es keine Angaben, sie finden demnach wohl nicht statt.

Brill und MDPI (Koninklijke Brill B.V. o.J.; MDPI AG 2024) sehen in ihren Datenschutzerklärungen keine Beschränkung auf die EU und Art. 45-Staaten vor.[21] Gleiches gilt für Elsevier, Frontiers und Wiley, bei denen zusätzlich eine nicht abschließende Liste von Drittstaaten genannt wird, darunter z.B. Indien, die VR China oder die Philippinen (Elsevier B.V. 2024; Frontiers Media SA 2023; John Wiley & Sons Inc. 2024). In solche Länder können Daten transferiert werden, wenn vertragliche Garantien der Verantwortlichen gegenüber dem Verlag vorliegen (Art. 46 DSGVO); Angemessenheitsbeschlüsse bestehen in Bezug auf diese Länder jedoch nicht. Damit kann aber kaum noch von einem dem europäischen Rechtsrahmen angemessenen Datenschutzniveau gesprochen werden. Bei Wiley besteht zusätzlich die Besonderheit, dass Vertragspartner der DEAL-Verträge die Wiley-VCH GmbH mit Sitz in Weinheim ist. Deren Datenschutzerklärung (Wiley-VCH GmbH 2022) ist an sich

wesentlich restriktiver als die des Mutterkonzerns. Da sie aber Datenübermittlung an andere Konzernteile vorsieht, gilt faktisch wohl doch die Datenschutzerklärung des Mutterkonzerns. Insgesamt ist zu konstatieren, dass Verlage mit Sitz in Deutschland darauf verzichten, Daten in Ländern zu verarbeiten, die datenschutzrechtlich besonders problematisch sind. Diese Position nehmen die Verlage entweder freiwillig ein oder – im Fall von Springer Nature – als Ergebnis der Vertragsverhandlung.

Mit Blick auf die *Datensparsamkeit* wurde versucht, die Verlage dazu zu bringen, die volle IP-Adresse der Nutzenden nur möglichst kurz zu speichern. Die IP-Adresse ermöglicht in vielen Fällen die eindeutige Identifikation einer Person und zählt deswegen zu den personenbezogenen Daten. IP-Adressen werden benötigt für COUNTER-Statistiken, standardisierte und von Bibliotheken viel genutzte Nutzungsstatistiken elektronischer Zeitschriften, Datenbanken und Bücher. Die zur Erstellung und Nachvollziehbarkeit der COUNTER-Statistiken möglicherweise längerfristig notwendige Speicherung der IP-Adressen könnte nach Löschung der beiden letzten Oktette bei IPv4[22] auch in anonymisierter Form erfolgen. Gemäß eigenen Angaben wird dies von Nomos praktiziert (Nomos Verlagsgesellschaft mbH & Co. KG 2022); Springer Nature hat sich hierzu im DEAL-Vertrag verpflichtet. Mit Elsevier und Wiley konnte diese Maßnahme dagegen nicht vereinbart werden, jedoch werden im ersten Vertragsjahr zumindest Workshops zur weiteren Erörterung des Sachverhalts durchgeführt, die bis Ende 2024 allerdings noch nicht stattgefunden haben.

7. Zusammenfassung und Bewertung

Wissenschaftstracking eröffnet – insbesondere auch beim Einsatz von Verfahren der Künstlichen Intelligenz – neue Geschäftsfelder für Verlage, die mit herkömmlicher Verlagstätigkeit nur noch wenig zu tun haben. Für die abwehrende Haltung der Verlage in den DEAL-Verhandlungen ist neben wirtschaftlichen

Interessen sehr wahrscheinlich auch bedeutend, dass Präzedenzfälle, z.B. mit Blick auf die gemeinsame Verantwortlichkeit, und die damit verbundenen Auswirkungen auf ihre Verträge mit anderen (internationalen) Wissenschaftseinrichtungen vermieden werden sollen.

Insgesamt bringen die Regelungen zum Datenschutz in den DEAL-Verträgen etwas Licht und – bei Elsevier und Wiley – zugleich auch viel Schatten: Einerseits gelang es erstmals in nationalen Verträgen, deutlich konkretere Bestimmungen zum Datenschutz zu vereinbaren. Andererseits ist mit Bedauern festzustellen, dass insbesondere beim Datentracking fast nur eine Wiederholung von gesetzlichen Vorgaben erfolgt. Insofern bringen die Verträge kaum Mehrwert gegenüber der geltenden Gesetzeslage. Im Übrigen stützen die Befunde zum Tracking mit Cookies nur begrenzt das „gute Kleinverlage / böse Großverlage"-Narrativ.

Aus den Erfahrungen der drei DEAL-Verhandlungen wie aus den Entwicklungen in Richtung immer detaillierterer und technisch-versierterer Datentracking-Techniken ist zu lernen, dass der Datenschutz frühzeitig Teil der Vertragsverhandlungen mit Verlagen werden muss und nicht erst nach der Einigung auf die wirtschaftlichen Rahmenbedingungen diskutiert werden sollte – dann ohne die Möglichkeit, Druck auszuüben. Eine Abstimmung der Wissenschaftseinrichtungen auf europäischer Ebene zur besseren Regulierung des Datentracking ist dringend vorzubereiten.

Es bleibt abzuwarten, ob die Verlage bereit sind, sich in einem partnerschaftlichen Dialog mit den Wissenschaften die Grundsätze der Transparenz und Datensparsamkeit zu eigen zu machen, die Datensouveränität der Wissenschaftler:innen wiederherzustellen und letztlich nicht weiter die Wissenschaftsfreiheit zu bedrohen (Kunz 2022).[23]

Disclaimer

Der Autor ist Mitglied der DEAL-Gruppe und war an DEAL-Verhandlungen beteiligt. Er war außerdem Mitglied der

AWBI/DEAL-Ad hoc-Arbeitsgruppe „Datentracking in den Wissenschaften“. Er ist Unterzeichner der Initiative „Stop Tracking Science“. Der Autor dankt Michaela Bilic-Merdes für die Durchsicht des Manuskripts.

Keine Aussage in diesem Beitrag soll als Behauptung interpretiert werde, Verlage würden gegen gesetzliche Bestimmungen verstoßen.

Anmerkungen

[1] Dieser Beitrag beruht auf dem Vortrag (Bilic-Merdes 2023) und den Aufsätzen (Altschaffel et al. 2024) und (Mittermaier 2024).

[2] https://www.textbroker.de/real-time-bidding.

[3] Viele hilfreiche Hinweise zur Datensparsamkeit im Alltag bietet das regelmäßig aktualisierte Buch (Zotzmann-Koch 2023).

[4] Die sog. COUNTER-Statistiken sind nach einem internationalen Standard. https://www.countermetrics.org/ erhobene Nutzungsstatistiken. Für Bibliotheken sind sie Grundlage zur Evaluierung der Nutzung elektronischer Lizenzen.

[5] https://www.elsevier.com/products/scopus/scopus-ai.

[6] Vgl. die Cookie Policy von Frontiers. https://www.frontiersin.org/legal/cookie-policy. Im Rahmen der Studie „Browser Fingerprinting.“ https://browser-fingerprint.cs.fau.de, wurde festgestellt, dass rund zwei Drittel der Nutzendenden eine einzigartige Browserkonfiguration haben und anhand dieser nachverfolgbar sind.

[7] Für Deutschland gilt der DEAL-Vertrag, der mit Wiley VCH GmbH, Weinheim, Deutschland abgeschlossen wurde. Deren Datenschutzbestimmungen https://www.wiley-vch.de/de/ueber-wiley/impressum#datenschutz, sehen solch weitgehendes Tracking nicht vor. Da die Nutzung aber faktisch auf den Seiten des US-amerikanischen Mutterkonzerns erfolgt darf bezweifelt werden, dass die restriktiveren Bestimmungen der deutschen Datenschutzbestimmungen eingehalten werden.

[8] Scripts sind Computerprogramme, die Bestandteil der jeweiligen Webseite sind und Befehle ausführen wie z.B. die Installation von Trackern.

[9] https://dsgvo-gesetz.de/art-5-dsgvo/.

[10] https://dsgvo-gesetz.de/erwaegungsgruende/nr-47/.

[11] Die Richtlinie 95/46/EG vom 24. Oktober 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr wurde 2016 von der DSGVO abgelöst. Artikel 7 der Richtlinie entspricht sinngemäß Artikel 6 der DSGVO.

[12] EuGH Urt. v. 5. Juni 2018, Rs. C210/16 – Facebook-Fan-Pages Rn. 36 f.; EuGH Urt. v. 10. Juli 2018, Rs. C25/17 – Zeugen Jehovas Rn. 73; EuGH, Urt. v. 29.7.2019, Az. C-40/17 – Fashion ID Rn. 80; EuGH Urt. v. 5. Dezember 2023, Rs. C683/21 – Nacionalinis visuomenes sveikatos centras Rn. 43 f.; EuGH Urt. v. 7. März 2024, Rs. C604/22 – IAB Europe Rn. 77.

[13] Bibliotheken übernehmen häufig, z.B. im Rahmen der DEAL-Verträge, Open Access-Publikationskosten für Autor:innen aus ihrer Einrichtung. Die Kostenübernahme setzt aber voraus, dass die Zugehörigkeit der Autor:innen zur Einrichtung im Rahmen eines Verifikationsprozesses bestätigt wird, vgl. z.B. https://bibblog.ub.uni-siegen.de/8206/neue-deal-vertraege-mit-elsevier-springer-nature-und-wiley/.

[14] https://stoptrackingscience.eu/.

[15] Die DEAL-Verhandlungen wurden kooperativ durch ein Verhandlungsteam unter Leitung von Horst Hippler geführt, damals Präsident der Hochschulrektorenkonferenz. Die Verhandlungsgruppe war aber keine Rechtsperson, die zur Vertragsunterzeichnung im Stande war. Mit Zustimmung der Gemeinsamen Wissenschaftskonferenz gründete die Max-Planck-Gesellschaft das Tochterunternehmen MPDL Services GmbH (MPDLS), die Vertragspartner der Verlage wurde. Inzwischen sind verschiedene Mitglieder der Allianz der Wissenschaftsorganisationen Gesellschafter der nun gemeinnützigen MPDLS gGmbH. https://mpdl-services.de/. Hiervon zu unterscheiden ist die MPDL (Max Planck Digital Library), die als Bibliothek der Max-Planck-Gesellschaft fungiert.

[16] Zu den Problemen mit diesem „Gummiparagraph" sei auf die Ausführungen in Abschnitt „Rechtlicher Rahmen für Datentracking" verwiesen.

[17] Datenschutzerklärung Nomos e-Library Stand 11.07.2022: „Temporäre Cookies (auch: Session- oder Sitzungs-Cookies): Temporäre Cookies werden spätestens gelöscht, nachdem ein Nutzer ein Online-Angebot verlassen und sein Endgerät (z.B. Browser oder mobile Applikation) geschlossen hat." https://www.nomos-elibrary.de/datenschutz.

[18] Der DEAL-Vertrag wurde mit Springer Nature Customer Service Center GmbH, Tiergartenstrasse 15–17, 69121 Heidelberg, geschlossen. Betreiber der Webseite, auf der der Zugriff auf die Publikationen erfolgt, ist allerdings Springer Nature Switzerland AG, Gewerbestrasse 11, 6330 Cham, Schweiz.

[19] Der DEAL-Vertrag wurde mit Wiley-VCH GmbH, Boschstraße 12, 69469 Weinheim geschlossen. Betreiber der Webseite, auf der der Zugriff auf die Publikationen erfolgt, ist allerdings John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030-5774, USA.

[20] https://commission.europa.eu/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en.

[21] Die EU-Kommission listet auf der Seite. https://commission.europa.eu/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en?prefLang=de, die bestehenden Angemessenheitsbeschlüsse auf, aktuell mit 15 Staaten.

[22] Die eindeutige Kennzeichnung eines Rechners besteht in der IPv4-Notation aus vier Zahlen („Oktette"), die Werte von 0 bis 255 annehmen können und mit einem Punkt getrennt werden, beispielsweise 134.94.59.79. Bei allen Rechnern des Forschungszentrums Jülich sind die beiden ersten Oktette („Class-B-Netz") identisch. Es würde deshalb eine Speicherung von 134.94.*.* genügen.

[23] Möglicherweise bringen die am 01.08.2024 gegen Nomos, Springer Nature und Wiley eingereichten Datenschutzbeschwerden zusätzliche Bewegung in die Angelegenheit (Gesellschaft für Freiheitsrechte, 2024).

Literatur

Altschaffel, R., Beurskens, M., Dittmann, J., Horstmann, W., Kiltz, S., Lauer, G., Ludwig, J., Mittermaier, B. & Stump, K. (2024) „Datentracking und DEAL: Zu den Verhandlungen 2022/2023 und den Folgen für die wissenschaftlichen Bibliotheken", *Recht und Zugang*, Vol. 5, No. 1, S. 23–40 [Online]. DOI: https://doi.org/10.5771/2699-1284-2024-1-23.

American Chemical Society (2024) *Privacy Policy* [Online]. Verfügbar unter https://www.acs.org/privacy.html (Abgerufen am 4 August 2024).

Artikel-29-Datenschutzgruppe (2014) *Stellungnahme 06/2014 zum Begriff des berechtigten Interesses des für die Verarbeitung Verantwortlichen gemäß Artikel 7 der Richtlinie 95/46/EG* [Online]. Verfügbar unter

https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_de.pdf (Abgerufen am 7 Juli 2024).

Augsberg, S.; Gehring, P. (2022) *Datensouveränität* [Online], Campus Verlag. Verfügbar unter https://doi.org/10.12907/978-3-593-45194-7.

Bettinger, E. C.; Bursic, M.; Chandler, A. (2023) *Disrupting the Digital Status Quo: Why and How to Staff for Privacy in Academic Libraries* [Online]. Verfügbar unter https://publish.illinois.edu/licensingprivacy/files/2023/06/Whitepaper-on-Privacy-Staffing-Licensing-Privacy.pdf (Abgerufen am 7 Juli 2024).

Beurskens, M. (2024) *Wissenschaftstracking. Rechtliche Einführung.* [Online]. Verfügbar unter https://www.dfg.de/resource/blob/329532/30b36867841f89efa6473cf30190bf6a/lis-awbi-folien-webinar-science-tracking-20240311-data.pdf (Abgerufen am 7 Juli 2024).

Bilic-Merdes, M. (2023) *Datentracking – Gefahr für die Wissenschaft?* Stuttgart.

Bilic-Merdes, M.; Lauer, G. (2023) *Datentracking – Gefahr für die Freiheit der Wissenschaft?* [Online]. Verfügbar unter https://dbt2023.abstractserver.com/program/#/details/presentations/424 (Abgerufen am 7 Juli 2024).

Burchard, A. (2013) „Nationallizenzen für Uni-Bibliotheken gefordert", *Tagesspiegel*, 11. August [Online]. Verfügbar unter https://web.archive.org/web/20130916131908/https://www.tagesspiegel.de/wissen/teure-fachzeitschriften-nationallizenzen-fuer-uni-bibliotheken-gefordert/8624114.html (Abgerufen am 15 August 2024).

DFG-Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme (2021) *Datentracking in der Wissenschaft: Aggregation und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage* [Online]. Verfügbar unter https://doi.org/10.5281/zenodo.5900759.

Elsevier B.V. (2024) *Privacy Policy* [Online]. Verfügbar unter https://www.elsevier.com/de-de/legal/privacy-policy (Abgerufen am 4 August 2024).

Elsevier B.V.; MPDL Services gGmbH (2023) *Projekt DEAL – Elsevier Publish and Read Agreement* [Online]. Verfügbar unter https://dx.doi.org/10.17617/2.3523659.

Fladung, R. B., Dugall, B. (2003) „Erwerbung. Ausweg aus der Zeitschriftenkrise? Ein Entscheidungsmodell für den Bezug elektronischer Zeitschriften im konsortialen Rahmen", *Bibliotheksdienst*, Vol. 37, No. 12, S. 1557–1574 [Online]. DOI: https://doi.org/10.1515/bd.2003.37.12.1557.

Franz Steiner Verlag (2024) *Datenschutzerklärung* [Online]. Verfügbar unter https://www.steiner-verlag.de/Service/Datenschutzerklaerung/ (Abgerufen am 4 August 2024).

Freiberg, M. (2022) „Third-Party-Tracking bei Wiley und Springer", *ABI Technik*, Vol. 42, No. 2, S. 96–104 [Online]. DOI: https://doi.org/10.1515/abitech-2022-0017.

Frontiers Media SA (2023) *Privacy Policy* [Online]. Verfügbar unter https://www.frontiersin.org/legal/privacy-policy (Abgerufen am 27 Juli 2024).

Georg Thieme Verlag KG (2021) *Datenschutzerklärung der Georg Thieme Verlag KG* [Online]. Verfügbar unter https://www.thieme.com/de-de/datenschutzerklaerung/georg-thieme-verlag-gmbh (Abgerufen am 4 August 2024).

Gesellschaft für Freiheitsrechte (2024) *Unzulässige Überwachung: Wissenschaftsverlage tracken die Aktivitäten von Forscher*innen* [Online]. Verfügbar unter https://freiheitsrechte.org/themen/demokratie/wissenschaftstracking (Abgerufen am 4 August 2024).

Hanson, C. (2019) *User Tracking on Academic Publisher Platforms* [Online]. Verfügbar unter https://www.codyh.com/writing/tracking.html (Abgerufen am 7 Juli 2024).

Hogrefe Verlag GmbH & Co. KG (2024) *Datenschutzerklärung Hogrefe Verlag GmbH & Co. KG* [Online]. Verfügbar unter https://www.hogrefe.com/de/datenschutzerklaerung (Abgerufen am 4 August 2024).

John Wiley & Sons Inc. (2024) *Privacy policy* [Online]. Verfügbar unter https://www.wiley.com/en-us/privacy (Abgerufen am 15 September 2024).

Kiltz, S.; Altschaffel, R.; Dittmann, J. (2023) „Science-Tracker Fingerprinting with Uncertainty: Selected Common Characteristics of Publishers from Network to Application Trackers on the Example of Web, App and Email", in Buchmann, E., Veith, E. & Cavendish, D. (Hg.) *SECURWARE 2023: The Seventeenth International Conference on Emerging Security Information, Systems and Technologies : September 25–29, 2023, Porto, Portugal* [Online], Wilmington, DE, USA, IARIA, 88–97. Verfügbar unter https://www.thinkmind.org/index.php?view=instance&instance=SECURWARE+2023 (Abgerufen am 7 Juli 2024).

Koninklijke Brill B.V. (o.J.) *Datenschutzerklärung* [Online]. Verfügbar unter https://brill.com/page/Privacy%20Statement/privacy-statement (Abgerufen am 4 August 2024).

Kunz, R. (2022) „Datentracking und die herausgeforderte Wissenschaftsfreiheit im digitalen Zeitalte", *Recht und Zugang*, Vol. 3, No. 1, S. 77–91 [Online]. DOI: https://doi.org/10.5771/2699-1284-2022-1-77.

Lamdan, S. (2022) *Data cartels: The companies that control and monopolize our information*, Stanford, California, Stanford University Press.

Lauer, G. (2022) „Datentracking in den Wissenschaften", *o-bib. Das offene Bibliotheksjournal*, Vol. 9, No. 1, S. 1–13 [Online]. DOI: https://doi.org/10.5282/O-BIB/5796.

Lindstrot, B. (2024) *Aufzeigen von Drittanbietern in den Webseiten von wissenschaftlichen Verlagen anhand von Codefragmenten im Seitenquelltext mit Fokus auf Datentracking*. Abschlussarbeit. TH Köln.

MDPI AG (2024) *MDPI Privacy Policy* [Online]. Verfügbar unter https://www.mdpi.com/about/privacy (Abgerufen am 4 August 2024).

Meier, M. (2002) *Returning science to the scientists: Der Umbruch im STM-Zeitschriftenmarkt unter Einfluss des Electronic Publishing* [Online], München, Peniope. Verfügbar unter http://webdoc.sub.gwdg.de/ebook/aw/2004/tel.ccsd.cnrs.fr/tel-00002257-00/tel-00002257.pdf (Abgerufen am 15 August 2024).

Mittermaier, B. (2023a) „DEAL: Wo stehen wir nach 10 Jahren? (Teil 1)", *b.i.t.online*, Vol. 26, No. 2, S. 123–131 [Online]. Verfügbar unter https://www.b-i-t-online.de/heft/2023-02-fachbeitrag-mittermaier.pdf (Abgerufen am 7 Juli 2024).

Mittermaier, B. (2023b) „DEAL: Wo stehen wir nach 10 Jahren? (Teil 2)", *b.i.t.online*, Vol. 26, No. 3, S. 217–225 [Online]. Verfügbar unter https://www.b-i-t-online.de/heft/2023-03-fachbeitrag-mittermaier.pdf (Abgerufen am 7 Juli 2024).

Mittermaier, B. (2024) „Datenschutz und Datentracking in Verlagsverträgen", *o-bib. Das offene Bibliotheksjournal*, Vol. 11, No. 4 [Online]. DOI: https://doi.org/10.5282/o-bib/6087.

Mittermaier, B.; Stanek, U. (2025) „Lizenzierung elektronischer Medien", in Johannsen, J., Mittermaier, B., Schäffler, H. & Söllner, K. (Hg.) *Praxishandbuch Bibliotheksmanagement* [Online], Berlin, Boston, De Gruyter Saur, S. 263–278. Verfügbar unter https://doi.org/10.1515/9783111046341.

Nomos Verlagsgesellschaft mbH & Co. KG (2022) *Datenschutzerklärung* [Online]. Verfügbar unter https://www.nomos-elibrary.de/datenschutz (Abgerufen am 4 August 2024).

OA2020 (2016) Expression of Interest in the Large-Scale Implementation of Open Access to Scholarly Journals. Online verfügbar unter https://oa2020.org/mission/, zuletzt aktualisiert am 21.03.2016, zuletzt geprüft am 07.07.2024.

Philosophy Documentation Center (o.J.) *Privacy* [Online]. Verfügbar unter https://www.pdcnet.org/wp/privacy/ (Abgerufen am 4 August 2024).

Pooley, J. (2022) „Surveillance Publishing", *Elephant in the Lab* [Online]. DOI: https://doi.org/10.5281/ZENODO.6384605.

Reuter, P.; Söllner, K. (2022) „Datentracking in der Wissenschaft – ein Sonderfall?", *ABI Technik*, Vol. 42, No. 2, S. 94–95 [Online]. DOI: https://doi.org/10.1515/abitech-2022-0016.

Roßnagel, A. (2022) „Third-Party-Tracking – ein Problem aus Sicht des Datenschutzes?", *ABI Technik*, Vol. 42, No. 2, S. 105–107 [Online]. DOI: https://doi.org/10.1515/abitech-2022-0018.

Schonfeld, R. C. (2018) *The Supercontinent of Scholarly Publishing?* [Online]. Verfügbar unter https://scholarlykitchen.sspnet.org/2018/05/03/supercontinent-scholarly-publishing (Abgerufen am 7 Juli 2024).

Schreiber, K. (2021) *Bezahlen mit Daten* [Online]. Verfügbar unter https://digitalisierungsrecht.eu/bezahlen-mit-daten/ (Abgerufen am 7 Juli 2024).

Siems, R. (2021) „When your journal reads you – user tracking on science publisher platforms", *Elephant in the Lab* [Online]. DOI: https://doi.org/10.5281/ZENODO.4683778.

Siems, R. (2022) „Das Lesen der Anderen", *o-bib. Das offene Bibliotheksjournal*, Vol. 9, No. 1, 1–25 [Online]. DOI: https://doi.org/10.5282/O-BIB/5797.

SPARC (2021) *Addressing the Alarming Systems of Surveillance Built By Library Vendors* [Online]. Verfügbar unter https://sparcopen.org/news/2021/addressing-the-alarming-systems-of-surveillance-built-by-library-vendors/ (Abgerufen am 7 Juli 2024).

Springer Nature Customer Service Center GmbH; MPDL Services gGmbH *Projekt DEAL – Springer Nature Publish and Read Agreement, 2024–2028* [Online]. Verfügbar unter https://doi.org/10.17617/2.3551270.

Springer Nature Group (2022) „Kurzstatement von Springer Nature", *ABI Technik*, Vol. 42, No. 2, S. 108 [Online]. DOI: https://doi.org/10.1515/abitech-2022-0019.

Springer Nature Switzerland AG (2024) *Privacy Statement* [Online]. Verfügbar unter https://link.springer.com/privacystatement (Abgerufen am 4 August 2024).

Vittorio Klostermann GmbH (2024) *Datenschutzerklärung* [Online] (Abgerufen am 4 August 2024).

Vogel, G. (2017) „German researchers start 2017 without Elsevier journals", *Science*, 355 (6320), S. 17 [Online]. DOI: https://doi.org/10.1126/science.355.6320.17.

Walter de Gruyter GmbH (2024) *Datenschutzerklärung* [Online]. Verfügbar unter https://www.degruyter.com/publishing/datenschutzerklaerung (Abgerufen am 4 August 2024).

Wiley-VCH GmbH (2022) *Impressum/Datenschutz* [Online]. Verfügbar unter https://www.wiley-vch.de/de/ueber-wiley/impressum (Abgerufen am 4 August 2022).

Wiley-VCH GmbH; MPDL Services gGmbH (2023) *Projekt DEAL – Wiley Publish and Access Agreement, 2024–2028* [Online]. Verfügbar unter https://dx.doi.org/10.17617/2.3551268.

Yoose, B.; Shockey, N. (2023) *Navigating Risk in Vendor Data Privacy Practices: An Analysis of Elsevier's ScienceDirect* [Online]. Verfügbar unter https://doi.org/10.5281/zenodo.10078610.

Yoose, B.; Shockey, N. (2024) *Navigating Risk in Vendor Data Privacy Practices: An Analysis of Springer Nature's SpringerLink* [Online]. Verfügbar unter https://doi.org/10.5281/zenodo.13886473.

Zotzmann-Koch, K. (2023) *Dann haben die halt meine Daten. Na und?!: Ein Buch für alle, die nichts zu verbergen haben*, Wien, edition silbenreich; edition sil!ben!reich.

KAPITEL 5

Transparenz statt Transzendenz: Orientierung im ABC der Publikationskosten

Lisa-Marie Stein

Publikationskosten sind Teil der Praxiswelt des Publizierens. Und auch beim elektronischen Publizieren entstehen Kosten. Subskriptionsbasierte und Open-Access-Publikationen unterliegen unterschiedlichen Geschäftsmodellen, Finanzierungsstrategien und Gebührenarten. Dieses Informationsgefüge in seiner Gesamtheit zu erfassen und zu begreifen gleicht mitunter auch für im Publikationsgeschäft Umtriebige einem Streben nach Transzendenz: Es liegt fern einer greifbaren Wirklichkeit. Transzendenz soll fortan durch Transparenz ersetzt werden. Anstelle eines Umgangs mit Publikationskosten, der umfangreiche Recherchen erfordert und über das unmittelbar Erfahr- oder Erklärbare hinausgeht, sollten diese sichtbar, klar strukturiert und überprüfbar sein. Eine adäquate Erfassung von Gebühren ermöglicht es, Kosten transparent nachvollziehen, zu vergleichen und für Verarbeitungszwecke

Zitationsvorschlag für dieses Kapitel:
Stein, L.-M. 2026. Transparenz statt Transzendenz: Orientierung im ABC der Publikationskosten. In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum*. S. 97–127. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.f. License: CC BY 4.0

automatisiert auszutauschen. In der wissenschaftlichen Infrastruktur gewinnen entsprechende Entwicklungen spürbar an Dynamik. Dennoch gilt es, weitere Hürden zu überwinden. Die im Zuge des Daten- bzw. Science Trackings[1] evozierte Transparenz der Wissenschaft – dort jedoch zu Ungunsten guter wissenschaftlicher Praxis – wird im Bereich der Publikationskosten ersehnt.

Der folgende Beitrag widmet sich dem Wandel und der Neustrukturierung, die u.a. die Open-Access-Transformation mit sich bringt. Das sogenannte Informationsbudget verlangt Prozessveränderungen in wissenschaftlichen Einrichtungen. Bibliotheken sind zunehmend in den Publikationsprozess involviert und entwickeln Change-Management-Strategien. Das Publikationskostenmanagement wird zu einer strategischen Herausforderung. Das Projekt *openCost*[2] nimmt sich dieser gezielt an und entwirft einen operativen Baustein zur systematischen Bewältigung der Anforderungen. Mit dem Ziel, Souveränität im Publikationsprozess zurückzugewinnen, soll durch Kostentransparenz nicht nur Fairness und finanzielle Planbarkeit, sondern insbesondere auch die wissenschaftliche Selbstbestimmung gestärkt werden, sodass Autor:innen informierte Entscheidungen treffen und im Rahmen der Publikationswahl Bedingungen vergleichen können. Sie werden damit zu aktiven Akteur:innen in der Transformation des Publikationswesens – nicht länger zu bloßen Konsument:innen oder Gebührenträger:innen degradiert.

Dazu ist zunächst eine systematische Beschreibung der relevanten Kostenarten und ihrer strukturellen Zusammenhänge erforderlich. Nach einer Skizzierung des Gegenstandsbereichs, der Finanzierungsmodelle und Gebühren, steht im folgenden Beitrag im Fokus, wie im Rahmen von openCost eine Offenlegung der vielfältigen Publikationskostenarten angestrebt wird. Es ergibt sich ein *Alphabet* der Kostenarten, das sich problemlos erweitern ließe.

Ziel des Projekts ist es, der Wissenschaft Übersicht und Kontrolle über diese Kosten zurückzugeben. In diesem Zuge entwickelte das openCost-Projektteam ein standardisiertes Metadatenschema für wissenschaftliche Einrichtungen, das diesen einen automatisierten

Austausch ihrer Kosteninformationen ermöglicht. Dieses und weitere Projektergebnisse sollen dazu beitragen, dass wissenschaftliche Institutionen bestehende Konstruktionen der Publikationsfinanzierung künftig mitverhandeln.

1. Publikationsverantwortung im Wandel der Finanzströme: Wissenschaftliche Bibliotheken & informierte Autorschaft

Wissenschaft lebt von einem möglichst ungehinderten Austausch von Ideen und Ergebnissen in Form von Publikationen. Diesen sicherzustellen ist seit jeher die Aufgabe von Bibliotheken. Neben dem Erwerb gedruckter Publikationen ist hierbei auch der Lizenzerwerb für Datenbanken und Zeitschriften seit Langem ein wesentlicher Bestandteil, und insbesondere Subskriptionsausgaben für Zeitschriften bilden häufig den mit Abstand größten Posten im Bibliotheksetat.

Publikationen spielen heute jedoch auch eine zentrale Rolle in der Leistungsbewertung: Sie sind praktisch die *Währung der Wissenschaft*. Publikationen führen zu Reputation und entsprechender finanzieller Ausstattung. Der Wert der Publikation wird hierbei häufig auf die Reputation des Publikationsorgans reduziert. Der so entstehende monopolistische Markt führte zu einem immensen Preisanstieg und in der Folge zu einer immer schlechteren Literaturversorgung.

Im Zuge der Digitalisierung und angestoßen durch die Open-Access-Bewegung befindet sich der Publikationsmarkt derzeit in einem Wandel von Closed Access hin zu einer wissenschaftsfreundlicheren Open-Access-Struktur. Im Rahmen der sogenannten Open-Access-Transformation sollen traditionelle Subskriptionszeitschriften in Open-Access-Zeitschriften umgewandelt und neue absolute Open-Access-Publikationsorgane etabliert werden.

Übergeordnetes Ziel ist ein freier und fairer Zugang zu wissenschaftlichen Ergebnissen und deren Rezeption. In der „Budapester Open Access Initiative“ (2002) wird dieser Zugang nicht nur

technisch sowie rechtlich als barrierearm beschrieben, sondern auch als für Lesende kostenfrei. Gleichzeitig wird erläutert, dass die Kostenfreiheit für Lesende nicht bedeute, dass im Publikationsprozess keine Kosten entstünden. Bereits 2002 wurde auf die Notwendigkeit zur Ausarbeitung neuer Kostendeckungsmodelle und Finanzierungsstrategien hingewiesen, auch wenn im Open Access bei steigender Reichweite insgesamt geringere Gesamtkosten als im traditionellen Subskriptionsmodell zu erwarten seien (ebd.).

Um diese Kosten zu erfassen und neuerlichen Fehlentwicklungen frühzeitig begegnen zu können, forderte der Wissenschaftsrat 2022 die Einführung eines sogenannten Informationsbudgets. Dieses umfasst sämtliche Kosten, die sowohl im Veröffentlichungsprozess als auch bei der Literaturbeschaffung verausgabt werden (vgl. Wissenschaftsrat 2022).[3] Notwendig ist hierzu eine vollständige Kostentransparenz und die Darstellung der Gesamtkosten einer Einrichtung, wie auch Heinz Pampel (2019: 11) im Hinblick auf die Herausforderungen der Open-Access-Transformation formuliert.

Die Bibliothek bietet sich für die Etablierung der hierzu nötigen zentralen Rechnungsbearbeitung für Publikationskosten an. Einerseits verwaltet sie bereits verschiedene Literaturbudgets, da schon im subskriptionsbasierten Modell in Bibliotheken die Verträge ausgehandelt und bearbeitet werden. Andererseits ist sie mit den Details des Publikationsprozesses vertraut und kann vertragsbedingte Zusammenhänge und Kostenübernahmen identifizieren, mit Autor:innen und Verlag unmittelbar in einen Informationsaustausch einsteigen oder fehlerhafte Rechnungen und unseriöse Predatory Journals erkennen – die zwar kostenintensiv, jedoch nicht qualitätssichernd agieren (vgl. open-access.network: „Predatory Publishing“).

Durch die Bündelung und den Einsatz jenes fachlichen Wissens an zentraler Stelle, die Etablierung wissenschaftsnaher Prozesse – die möglichst viele Anforderungen gleichzeitig bedienen – und eine zentrale (Meta-)Datenerfassung werden Wissenschaftler:innen und andere Bereiche von publikationsspezifischem bürokratischem

Aufwand entlastet. Neben der individuellen Beratung entwickeln und nutzen Bibliotheken hierbei Recherchewerkzeuge und -plattformen, wie die *Elektronische Zeitschriftenbibliothek*[4] (vgl. EZB), den *Open policy finder* (vgl. Open policy finder) oder das *Directory of Open Access Journals*[5] (vgl. DOAJ), um zu Publikationswegen, Lizenzen und Verlagsbedingungen zu beraten.

Das zentrale Kostenmonitoring und die -verwaltung erfordern jedoch eine intensive Zusammenarbeit bislang häufig getrennter Abteilungen innerhalb einer wissenschaftlichen Einrichtung und gehen mit entsprechenden Umstrukturierungen einher. Die Zuständigkeitsbereiche für Lesezugriffe und Veröffentlichungsprozesse rücken enger zusammen. Aus Sicht der Nutzer:innen tritt neben einer optimalen Informationsversorgung auch eine zielgerichtete Publikationsberatung in den Vordergrund, die es ihnen ermöglicht, informierte und reflektierte Entscheidungen über ihre Publikationswege zu treffen. Durch nachvollziehbare Angaben zu Gebühren und Finanzierungsmodellen werden Verlagsbedingungen vergleichbar.

Die *Deutsche Forschungsgemeinschaft*[6] (DFG) unterstützt mit dem Programm „Open-Access-Publikationskosten" die notwendige Umstrukturierung zur transparenten Erfassung und Zentralisierung der Finanzströme. Neben der Kofinanzierung der Open-Access-Publikationen einer Einrichtung sind alle an diesem Programm teilnehmenden Einrichtungen zu einem detaillierten Monitoring ihrer Publikationskosten und zentralen Ablieferung entsprechender Berichte verpflichtet.

Neben der DFG gibt es noch weitere Interessenten an Kostendaten, wie andere Forschungsfördernde oder Informationsplattformen (vgl. z.B. *OpenAPC*) und Datenbanken (vgl. z.B. *OpenAlex*). Um die Informationsaufbereitung nicht für jeden Anlass individuell anfertigen zu müssen, entwickelt das ebenfalls DFG-finanzierte Projekt openCost eine standardisierte und durch Etablierung einer Schnittstelle automatisierte Lösung, auf die nachfolgend näher eingegangen wird.

Die Open-Access-Transformation brachte bisher zahlreiche verschiedene Finanzierungsmodelle und Vertragskonstruktionen

hervor. Die Details der Publikationsworkflows und die kursierenden Definitionen sind komplex, ihre Handhabung bedarf fachlicher Expertise. Seit den ersten Verträgen haben sich Mechanismen fortentwickelt, Modelle wurden ausgearbeitet und neue Verträge verhandelt. Um die Tragweite der tatsächlich anfallenden Kosten diskutieren zu können, müssen diese Mechanismen offengelegt werden und für die beteiligten Akteure einsehbar sein. Publikationsberatung und -(kosten)management werden somit im digitalen Zeitalter zur zentralen Aufgabe wissenschaftlicher Bibliotheken. Der wissenschaftliche Publikationsprozess muss strukturiert, Geschäfts- und Finanzierungsmodelle nachvollziehbar und darunter auch Publikationskostentypen identifizierbar werden (vgl. dazu auch Deinzer/Lang 2024 zur Etablierung des Informationsbudgets an der Universität Regensburg).

Gleichzeitig bleibt in der Praxis die Herleitung der Preisgestaltung im Publikationsmarkt jedoch oftmals intransparent. Wie kommt die Gesamtsumme einer Vertragszahlung zustande? Was sind die Kosten für eine Publikation? Wie errechnen sich Publish- und Read-Anteile? Das Publikationsmanagement einer Einrichtung muss sich diesen Fragen stellen und des Weiteren mit vager Terminologie oder Informationsarmut auf Rechnungen umgehen: Nicht selten fehlen Identifikatoren, die eine Rechnung eindeutig einer konkreten Publikation zuweisen. Auch die Open-Access-Optionen *grün, gold* oder *diamond* sind zwar eine weit verbreitete Kategorisierung der Publikationswege, doch über ihre abgrenzbaren Definitionen herrscht nicht zwangsläufig Konsens. Beschreiben die auf einer konkreten Rechnung genannten Article Processing Charges eine Publikation in einer Open-Access-Zeitschrift oder Gebühren zum Freikauf eines Artikels in einem Closed-Access-Journal – im Zeitschriftenbereich sogenanntes hybrides Open Access – und ab wann gilt ein Vorgang überhaupt als hybrid?[7]

Grundlage zur Beantwortung dieser Fragen ist die Erfassung und Aufzeichnung von Kosten und Bedingungen. Nur so ist ein Überblick über die Ausgaben des Publizierens möglich. Ein freier Austausch der erfassten Daten erlaubt ferner nicht nur eine einrichtungsindividuelle Kontrolle, sondern auch Vergleiche

zwischen verschiedenen Institutionen. Auch die Effizienz konsortialer oder nationaler Verträge lässt sich nur mit Hilfe der Gesamtzahlen aller beteiligten Einrichtungen evaluieren. Erfassung, Aufzeichnung und Austausch der Kostendaten sind somit ein essentieller Baustein auf dem Weg zur Kostentransparenz.

Und diese ist schließlich Voraussetzung, um sich bewusst für oder gegen bestimmte Publikationsangebote entscheiden zu können. Dies stärkt die Autonomie der Forschenden, die sich damit aus einer Abhängigkeit von kommerziellen Anbietern und undurchsichtigen Preisstrukturen befreien können. Gerade im Zeitalter digitaler Plattformen gewinnt die informationelle Selbstbestimmung an Bedeutung, denn immer häufiger beeinflussen ökonomische Interessen, welche Forschung sichtbar gemacht wird – und welche nicht. Transparente Kostenmodelle geben Wissenschaftler:innen nicht nur Souveränität über ihre Publikationswahl zurück, sondern ermöglichen auch eine aktive Mitgestaltung des wissenschaftlichen Kommunikationssystems. Wer über konkrete Daten verfügt, kann sich fundiert in Debatten einbringen, z.B. zur OpenAccess-Transformation oder zur Gestaltung von Förderstrukturen. Darüber hinaus verbessert Transparenz die finanzielle Planbarkeit: Publikationskosten werden kalkulierbar, was dem Informationsbudget einer Einrichtung, aber besonders auch Forschenden mit wenig finanziellem Spielraum oder ohne institutionelle Unterstützung zugutekommt. Langfristig trägt Kostentransparenz somit auch zu mehr Fairness im Wissenschaftssystem bei. Sie wirkt strukturellen Ungleichheiten entgegen, indem sie die Voraussetzungen schafft, unter denen Zugang zu Wissen und die Veröffentlichung eigener Forschung unabhängig von Fachrichtung, Herkunft oder finanziellen Ressourcen möglich werden.

2. Publikationskosten: Transparenz mit openCost

Das durch die DFG geförderte Projekt openCost[8] (vgl. GEPRIS 2024) wurde im Jahr 2021 durch die Universitätsbibliotheken

Bielefeld und Regensburg sowie das Hamburger Forschungszentrum Deutsches Elektronen-Synchrotron DESY ins Leben gerufen. Der Projektuntertitel „automatisierte, standardisierte Lieferung und offene Bereitstellung von Publikationskosten und Verlagsvereinbarungen" beschreibt bereits das Programm. Stets unter der Prämisse der *FAIR*ness (Wilkinson 2016: 4) und Kostentransparenz ist die Offenlegung aller Publikationskosten in wissenschaftlichen Institutionen das Ziel. Dazu entwirft das Projekt eine technische Infrastruktur, mit der Publikationskosten in einem Standardformat erfasst und durch standardisierte Schnittstellen offen zugänglich sowie automatisiert bereitgestellt, abgefragt und ausgetauscht werden können. Die Projektschwerpunkte lassen sich wie folgt skizzieren (vgl. auch openCost-Webseite 2024):

(a) Ein standardisiertes Metadatenschema für Publikationskosten wird ausgearbeitet (vgl. openCost-Github-Repository).
(b) Für das Konsultieren der Publikationskostendaten und den automatischen Austausch muss eine offene Schnittstelle implementiert werden: Das *Open Archives Initiative Protocol for Metadata Harvesting*[9] (vgl. OAI-PMH 2024) findet hier Anwendung.
(c) Es erfolgt eine Erweiterung des Datenmodells und der Services der *EZB* (vgl. EZB 2024): Funktionen zur Anzeige, Verwaltung und Übernahme von Publikationskosten werden ergänzt.

Nachfolgend sollen die beim elektronischen Publizieren anfallenden bzw. von Publikationsorganen benannten Kostentypen genauer betrachtet werden, wie diese im openCost-Schema als Metadateninformation dargelegt (vgl. (a)) und schließlich systematisch, standardisiert und automatisiert zur weiteren Verarbeitung zur Verfügung gestellt werden können (vgl. (b)).

2.1 Das Alphabet der Publikationskosten

Anfang der 1990er Jahre entstand mit *arXiv* (vgl. arXiv 2024) ein Repositorium im Fachbereich der Physik, das zunächst als

Emailserver startete, aber nach Entstehung des World Wide Web zum Webserver fortentwickelt wurde (Drößler 2022: 82).

Insbesondere in der Teilchenphysik hat sich damit die Praxis etabliert, wissenschaftliche Ausarbeitungen unmittelbar offen zugänglich zu machen – im Sinne des Repositorienprinzips – und sie zunächst als sogenannte Preprints zu veröffentlichen. In vielen anderen akademischen Disziplinen wird dagegen eine Veröffentlichung ohne ein offizielles Peer-Review-Verfahren zur Qualitätssicherung, das gegebenenfalls bei der Manuskriptversion noch nicht stattgefunden hat, abgelehnt (ebd.: 83).

Auch in Bezug auf das Publizieren in Zeitschriften oder etwa die Ausarbeitung von Monographien bestehen fachspezifische Unterschiede: In unterschiedlichen Fachdisziplinen zeigen sich jeweils eigene Publikationskulturen. Welche Implikationen ergeben sich daraus für die detaillierte Beschreibung von Publikationskostentypen? Es treten nicht nur (Kap. 2.1.1) unterschiedliche Strategien und Strukturen sowie Geschäfts- und Finanzierungsmodelle zutage, sondern auch (Kap. 2.1.2) verschiedene Publikationsobjekte der jeweiligen wissenschaftlichen Disziplinen. Beide Aspekte bedingen die Ausgestaltung der Kostentypen (Kap. 2.1.3).

2.1.1 Geschäfts- und Finanzierungsmodelle

Soll eine Publikation frei zugänglich veröffentlicht werden, kann der Verlag Gebühren verlangen. Handelt es sich um eine Veröffentlichung in einer originären Open-Access-Zeitschrift (*Goldenes Open Access*), wird in Article-Processing-Charge-basierten[10] Modellen ein konkreter Preis eines Einzelartikels erhoben. Erfolgt die Publikation in einem Closed-Access-Journal, wird der offene Lesezugriff für den einzelnen Artikel erkauft und die Rede ist von *hybridem Open Access*.

Neben Open-Access-Gebühren können weitere Kosten anfallen, z.B. eine Einreichungsgebühr oder die Finanzierung eines Lektorats. Der Verlag kann des Weiteren mit einer wissenschaftlichen

Einrichtung in einem Vertragsverhältnis stehen, das spezifische und individuelle Konditionen zur Finanzierung des Publikationsprozesses für alle oder einige seiner Zeitschriften aushandelt. So können je nach Zeitschriftenportfolio sowohl goldene als auch hybride Publikationskosten sowie gegebenenfalls anfallende Zusatzgebühren im Rahmen eines Vertrags beglichen werden. Ein sogenannter *Transformationsvertrag* soll Zeitschriften den Übergang vom subskriptionsbasierten Abonnement zum Open-Access-Journal erleichtern.

Fest steht jedoch, dass sowohl subskriptionsbasierte als auch Open-Access-Verträge finanzorientiert aufbereitet werden müssen. Unter letzteren, um die es hier primär gehen soll, gibt es wiederum unterschiedliche Herangehensweisen und Vertragsinhalte, die häufig in konsortialen Zusammenschlüssen verhandelt werden. Offsetting-Modelle, Publish & Read-Verträge, Gold-Open-Access-Verträge, Crowdfunding oder Konzepte wie *Subscribe to Open* und *Purchase to Open* sind nur einige transformative Szenarien. Es existieren unterschiedliche Herleitungs- und Finanzierungsstrategien, deren Kosten wiederum entsprechend formal abgebildet werden müssen.

Ob und zu welchem Zeitpunkt die Open-Access-Gebühr einer Einzelpublikation festgestellt werden kann, ist wieder durch eine potentielle Vertragszugehörigkeit oder eine davon unabhängige Finanzierung bedingt. So wird beispielsweise der Artikelpreis von Gold-Open-Access-Publikationen im Rahmen von Transformationsverträgen explizit auf Einzel- oder Sammelrechnungen ausgewiesen. Für hybride Veröffentlichungen unter einem transformativen Vertrag, die zur Open-Access-Publikation freigekauft werden, ist die mitunter intendierte Herleitung einer effektiven APC ein komplexeres Unterfangen. Die Vertragsmodelle basieren auf unterschiedlichen Berechnungslogiken. So wird etwa eine bisherige Subskriptionszahlung durch eine Publikationszahlung ersetzt, während zugleich APC-Beträge pro Publikation berechnet werden. Diese erweisen sich jedoch rückblickend – gemessen am tatsächlichen Publikationsaufkommen – häufig als wenig belastbare Werte.

Die fehlende Verlässlichkeit ergibt sich aus einer Vielzahl von Faktoren: Dazu zählen unter anderem Schätzungen zum zukünftigen Publikationsaufkommen, die Einbeziehung von Lesezugriffen, finanzielle Zusammenschlüsse wissenschaftlicher Einrichtungen sowie Rabatte, Sonderkonditionen und nachträgliche Ausgleichszahlungen. Vor diesem Hintergrund ist eine präzise Bestimmung des tatsächlichen Artikelpreises zum Zeitpunkt der Veröffentlichung faktisch nicht möglich.

Diese strukturelle Intransparenz erschwert es insbesondere Wissenschaftler:innen, die tatsächlichen Kosten ihrer Open-Access-Publikationen nachzuvollziehen und fundierte Entscheidungen über Publikationswege zu treffen.

Auch für nicht-kommerzielle und wissenschaftsgeleitete Diamond-Open-Access-Modelle (Deutsche Forschungsgemeinschaft 2024: 2) bedarf es einer Abbildung der individuellen Kostenstrukturen. Pampel et al. formulieren „Thesen zur Zukunft des wissenschaftsgeleiteten Open-Access-Publizierens […]", die bei der Satelliten-Konferenz „Wissenschaftsgeleitetes Open-Access-Publizieren" 2023 am Institut für Bibliotheks- und Informationswissenschaft der Humboldt-Universität zu Berlin diskutiert und anschließend aufgearbeitet wurden (Pampel et al. 2024: 322). Mit ihrer These „Diversität der Finanzierungs- und Geschäftsmodelle anerkennen" unterstreichen sie nicht nur die Notwendigkeit zur Kostentransparenz, sondern auch den Bedarf der Berücksichtigung von Bibliodiversität: „Wissenschaftsgeleitete Publikationsinfrastrukturen bedürfen je nach (fachlicher) Community unterschiedlicher Finanzierungs- und Geschäftsmodelle, um den Anforderungen der jeweiligen (Fach-)Communities zu entsprechen. Dabei sollen größtmögliche Kosten- und Leistungstransparenz leitend sein" (Pampel et al. 2024: 330). Jene Diversität ist entsprechend mit den nachfolgend diskutierten Publikationsobjekten (Kap. 2.1.2) und -kostentypen (Kap. 2.1.3) abzubilden.

So existieren diverse Modelle und Finanzierungsstrategien, die eine transparente Dokumentation der Kosten zur Herausforderung machen. Transparenz würde durch eine strategische Modellierung der Geschäftsgänge wissenschaftlicher Einrichtungen

begünstigt, indem wie in Kapitel 1 beschrieben alle Publikationskosten durch die Bibliothek zentral erfasst werden. Der Transparenz hinderlich sind wiederum insbesondere jene Verträge, bei denen Ausgaben nicht nachvollziehbar bzw. nicht auf einen Artikelpreis umgerechnet werden können, wie auch Frick & Kaier (2020: 3) postulieren. Sie unterscheiden in diesem Zusammenhang zwischen „Publikationsgebühren" und „Pauschalen". Erstere beziehen sich auf den konkreten Preis eines Artikels, letztere werden häufig durch Vorauszahlungen realisiert, und der Artikelpreis kann nicht oder erst in der Retrospektive ermittelt werden (ebd.: 4).

2.1.2 Das publizierte Objekt

Die bisherigen Ausführungen, z.B. zu goldenen und hybriden Zeitschriften, beschreiben einen Publikationskontext, der relativ stark an der Publikationsform eines Zeitschriftenartikels orientiert ist. Der wissenschaftliche Zeitschriftenartikel spielt wohl unbestritten in der Open-Access-Transformation eine große Rolle. Nichtsdestotrotz werden wissenschaftliche Ergebnisse in Form unterschiedlicher Objekte produziert. Daher gilt es, das jeweilige Publikationsobjekt als kostentragende Entität samt seiner im Publikationsprozess erforderlichen Kostenpunkte zu berücksichtigen. In Tabelle 5.1 werden einige Beispiele jener Publikationsentitäten aufgeführt.

Um die Open-Access-Freischaltung eines Zeitschriftenartikels für den Verlag finanziell tragfähig zu machen und die im Publikationsprozess anfallenden Kosten abzudecken, berechnet dieser sogenannte *Article Processing Charges*, die unter der Abkürzung *APC* bekannt sind. APCs finden sowohl für die Open-Access-Publikation in vollständigen Open-Access-Zeitschriften als auch in sogenannten hybriden Zeitschriften Anwendung.

Neben einem permanenten offenen Zugriff auf den Artikel decke eine APC alle Kosten, die in den jeweiligen Phasen des Veröffentlichungsprozesses anfielen, von der Verwaltung des Peer-Reviews über das Lektorat bis hin zum Hosting des fertigen Artikels auf speziellen Servern, definieren die Verlage „Springer" und „Wiley"

und beschreiben damit einige kostentragende Tätigkeiten, die in den dortigen Prozessen in einer APC enthalten sind (vgl. Springer Support 2024 & Wiley Authorservices 2024). Qualitätssicherung, Veröffentlichung und die digitale Bereitstellung sind demnach leitend. „Processing" inkludiert bei Springer ferner redaktionelle Tätigkeiten (Peer-Review, administrative Unterstützung, Auftragsvergabe für Inhalte und die Entwicklung der Zeitschrift), technische Infrastruktur und Innovation (Entwicklung, Wartung und Betrieb von Online-Zeitschriftensystemen und Websites), die Aufbereitung der Artikel (Formatierung und Markierung von Artikeln sowie Aufnahme in Indexierungsdienste), die Vermarktung der Zeitschrift und der Inhalte sowie Kundendienste (Beantwortung von Anfragen von Autor:innen und Leser:innen) (vgl. ebd.).

In der Regel werden APCs zum Zeitpunkt der Annahme des Beitrags zur Veröffentlichung berechnet und die Kosten von dem/der Autor:in bzw. dessen Einrichtung getragen (vgl. open-access.network: „APC").

Nach gleichem Prinzip werden Kosten für Bücher (Book Processing Charges) und einzelne Kapitel (Chapter Processing Charges) fällig. Analoge Gebühren können z.B. für Forschungsdaten (Data Processing Charges), Filme bzw. audiovisuelle Medien (Film Processing Charges) oder Software (Software Processing Charges) erforderlich werden. Obgleich diese Fälle bislang noch weniger bekannt sind, ist auch hier anzunehmen, dass analoge Bedarfe, wie die wissenschaftliche Qualitätsfeststellung, der Veröffentlichungsprozess oder eine langfristige Verfügbarmachung sowie Archivierung, kostendeckend berechnet werden. Bereits jetzt bieten Verlage mitunter an, Forschungsdaten begleitend zum publizierten Artikel zu veröffentlichen und eine gegenseitige Referenzierbarkeit sicherzustellen. Das Publizieren der Forschungsdaten in Repositorien ist zweifellos von steigender Relevanz.

Zusammenfassend besteht hier also die Frage nach dem zu publizierenden Objekt, das insbesondere auch nach Fachdisziplin variiert, während einige, wie z.B. der Zeitschriftenartikel, das Buch oder die Forschungsdaten, allgegenwärtig sind.

Aus dieser Analogie entspringt die Idee eines Alphabets für Publikationskosten (vgl. Tabelle 5.1). In diesem sind die Termini A bis

Tabelle 5.1: Publikationskosten-ABC.

Bezeichnung:		**Definition:**
APC (1)	Article Processing Charge	Gebühr zur Publikation eines Artikels
BPC (1)	Book Processing Charge	Gebühr zur Publikation eines Buches
CPC (1)	Chapter Processing Charge	Gebühr zur Publikation des Kapitels eines Buches
DPC (1)	Data Processing Charge	Gebühr zur Publikation von Forschungsdaten
EPC (2)[11]	Electronic Publishing Contracts	Oberbegriff für im Kontext eines Vertrags verzeichnete Publikationsgebühren
FPC (1)	Film Processing Charge	Gebühr zur Publikation von Filmmaterial/audio-visuellen Medien
GPC (1)	Graphic Processing Charge	Gebühr zur Publikation von Grafiken/Bildmaterial
…		
IPC (2)	Infrastructure Publishing Costs	Oberbegriff für Zahlungen einer Open-Science-Infrastruktur, die das wissenschaftliche Publizieren bedingen
…		
RPC (1)	Recording Processing Charge	Gebühr zur Publikation von Audioaufnahmen
SPC (1)	Software Processing Charge	Gebühr zur Publikation von Software
…		

C (APC, BPC und CPC) in der Praxis belegt, insbesondere die APC ist fest etabliert. Alle weiteren Gebühren beschreiben belegte Kostenpunkte, die gewählten Akronyme haben jedoch derzeit noch keine spezifische Referenz auf Publikationsrechnungen. Das Alphabet zeigt hierbei das Spektrum der Kosten, die Tendenz zur

Variation und v.a. die Bandbreite an Bereichen, in denen künftig Kosten im *Wissenschaftsmarkt* veranschlagt werden könnten.

Exemplarisch sei hier der Bereich der Forschungsdaten herausgegriffen (vgl. z.B. die verschiedenen *NFDI-Sektionen, -Konsortien* und *BASE4NFDI* (vgl. NFDI-Webseite)), die *Helmholtz Metadata Collaboration* (vgl. HMC-Webseite & Wagner et al. 2023) oder das *Open Science Festival* 2024 (vgl. Wagner 2024). Vor dem Hintergrund der *FAIR*-Prinzipien werden hier vielerlei infrastrukturelle Innovationen sichtbar. So werden nicht nur die Kriterien zum Status einer wissenschaftlichen Publikation zu definieren sein, sondern auch ihre finanziellen Bedarfe, die u.a. – aber nicht nur – aus der Notwendigkeit großer Speicherkapazitäten hervorgehen. Auch die Kosten einer Forschungsdatenpublikation sollten nachvollziehbar und strukturiert erfasst werden.

Innerhalb der Gebührenkategorie für Forschungsdaten (DPC) nimmt das publizierte Objekt wiederum diverse mediale Formen an. Das Format wird nicht nur durch den disziplinspezifischen Gegenstandsbereich bedingt, sondern auch durch seine Repräsentation zum jeweiligen Forschungszeitpunkt, sodass unterschiedliche Anforderungen und Versionen bedeutsam werden: „Als Forschungsdaten werden alle Daten bezeichnet, die während des Forschungsprozesses entstehen oder sein Ergebnis sind. Sie werden abhängig von der Forschungsfrage und unter Anwendung verschiedener Methoden erzeugt bzw. gesammelt, bearbeitet, analysiert und schließlich publiziert und/oder archiviert. Demzufolge treten die Forschungsdaten in jeder Wissenschaftsdisziplin in unterschiedlichen Medientypen, Aggregationsstufen und Formaten auf“ (Akademischer Senat der Humboldt-Universität zu Berlin 2014: 1).

Das *Publikationskosten-ABC* (Tabelle 5.1) benennt Open-Access-Gebühren, die terminologisch das publizierte Objekt beschreiben. Abweichend davon beziehen sich die EPCs (Electronic Publishing Contracts) auf die Beschreibung eines Vertragsobjekts im Kontext des elektronischen Publizierens, wie z.B. Transformations- oder Publish & Read-Verträge und Mitgliedschaften bei reinen

Open-Access-Verlagen. Sie subsumieren alle im Rahmen eines Vertrags anfallenden Kostentypen. EPCs definieren demnach ein kostentragendes Objekt im Rahmen des wissenschaftlichen Publikationswesens, nicht aber das publizierte Objekt.

Gleiches gilt für den Oberbegriff IPC (Infrastructure Publication Costs), der die infrastrukturellen Finanzströme von Open-Science-Akteuren dokumentiert. IPC benennt Zahlungen für Strukturen, die nicht unmittelbar aus dem Publikationsprozess selbst resultieren, sondern diesen begleiten, thematisieren oder unterstützen. Auch hier können Vertragsstrukturen vorliegen. Die *Global Sustainability Coalition for Open Science Services (SCOSS)* definiert Open-Science-Infrastrukturen als Ressourcen und Dienste für die wissenschaftliche Kommunikation, die von nicht-kommerziellen Organisationen bereitgestellt werden und es der wissenschaftlichen und akademischen Gemeinschaft ermöglichen, Forschungsergebnisse zu sammeln, zu speichern, zu organisieren, zugänglich zu machen, gemeinsam zu nutzen und zu bewerten (vgl. SCOSS 2024).

Graphic Processing Charges (GPC) bezeichnen nicht jene Zusatzkosten, die etwa für Abbildungen innerhalb eines Textbeitrags oder auf dem Cover eines Buchs oder einer Zeitschrift anfallen, sondern beziehen sich auf grafische Hauptobjekte, die selbst den Kern der Veröffentlichung darstellen und unabhängig von einer begleitenden Textpublikation einen freien Zugang erhalten sollen.

2.1.3 Publikationskostentypen

Verschiedene Geschäfts- und Finanzierungsmodelle sowie mediale Diversität bringen unterschiedliche Kostenarten hervor. Strukturell- (Kap. 2.1.1) und objektbedingt (Kap. 2.1.2) existieren also Publikationskostentypen, die im Folgenden genauer definiert werden sollen.

Die Publikationskostentypen (Tabellen 5.2 und 5.3) stellen keine nebengeordnete Erweiterung des Publikationskosten-ABC dar, sondern beziehen sich auf das übergeordnete Objekt (z.B. *Article, Book, Chapter,* …), das veröffentlicht werden soll bzw. für das Kosten verausgabt werden. In der folgenden Übersicht werden Kosten aufgeführt, die für eine Publikation anfallen können:

Tabelle 5.2[12]: Cost Type Publication.

Bezeichnung:	**Definition:**
gold-oa	Gebühren für den Open-Access-Status in einem reinen Open-Access-Publikationsorgan. Terminologisch i.d.R. auf einer Rechnung in Abhängigkeit des publizierten Objekts benannt als *APC, BPC*, etc.
hybrid-oa	Gebühren für den Open-Access-Status in einem Closed-Access-Publikationsorgan. Terminologisch i.d.R. auf einer Rechnung in Abhängigkeit des publizierten Objekts benannt als *APC, BPC*, etc.
vat	Mehrwertsteuer (value added tax)
colour charge	Gebühren für den Farbdruck (farbige Abbildungen, Seiten, Illustrationen etc.), Relevanz bei Printversionen
cover charge	Gebühren zur Abbildung eines Covers (Publikation von Inhalten zur Umschlag-/ Deckblattgestaltung)
page charge	Gebühren für die Überlänge oder eine seitenbasierte Abrechnung (letztere erscheint ggf. nicht als Zusatzgebühr, sondern ersetzt funktional eine Processing Charge, den Bereichen gold- oder hybrid-oa zuzuordnen)
permission	Gebühren für den Erwerb einer Lizenz (Nutzungsrecht, z.B. für eine Abbildung, um diese innerhalb einer Publikation nutzen zu dürfen)
publication charge	Publikationsgebühren, die im Rahmen von Closed Access zu leisten sind
reprint	Nachdruckgebühren für Kopien von Veröffentlichungen
submission fee	Gebühren zur Einreichung einer Publikation
payment fee	Gebühren zur finanziellen Transaktion der Zahlung (z.B. Bank-, Transfer- oder Kreditkartengebühren)

(*Continued*)

Tabelle 5.2[12]: Continued.

Bezeichnung:		**Definition:**
other		sonstige Gebühren
	abstract charge	Gebühren zur Einreichung oder Veröffentlichung des Abstracts (Zusammenfassung/Kurzreferat) einer Publikation (häufig im Kontext von Konferenzen und Tagungsbänden)
	archiving charge	Gebühren für die Archivierung einer Publikation
	editing charge	Gebühren für die Redaktion und das wissenschaftliche Lektorat
	licence fee	Gebühren für die Zuschreibung einer Lizenz zur Festlegung der Nutzungsrechte durch Dritte (zusätzlich zur gesamten Veröffentlichung ggf. auch zu bestimmen für enthaltene Bilder, Abdrucke, Tabellen, Grafiken etc.)
	print charge	Gebühren für den Druck einer Publikation
	translation charge	Gebühren für die Übersetzung einer Publikation
…		

Tabelle 5.3 beinhaltet die Kostentypen, die gemäß der in Tabelle 5.1 mit (2) markierten kostentragenden Entitäten Kostenkontexte beschreiben. Es findet keine artikelbasierte, sondern eine übergeordnete Zahlung statt, wie es bei vielen Vertragsabschlüssen der Fall ist.

2.2 Kosten als Metadateninformation der Publikation

Die in Abschnitt 2.1 ausgeführten Definitionen dienen nicht nur der theoretischen Orientierung im Dschungel wissenschaftlicher Publikationsmodalitäten, sondern auch der ganz praktischen Nutzung.

So initiierte openCost in Kooperation mit der Universitätsbibliothek der FAU Erlangen-Nürnberg ein Publikationskostenglossar,

Tabelle 5.3[13]: Cost Type Contract.

Bezeichnung:		**Definition:**
publish		Gebühren für das Open-Access-Publizieren im Rahmen von (z.B. *Publish & Read*-) Verträgen, Pauschale zur Deckung aller Publikationsdienstleistungen für das Veröffentlichen in unter den Vertrag fallenden Publikationsorganen
read		Gebühren für den Lesezugriff eines Closed-Access-Portfolios im Rahmen von (z.B. *Publish & Read*- oder *Subskriptions*-) Verträgen
vat		Mehrwertsteuer (value added tax)
other		
	service fee	Zusätzliche Gebühren zur Deckung von Service- und Betriebskosten
	infrastructure charge	Kostenbeitrag zur Finanzierung einer wissenschaftsgeleiteten Infrastruktur

das über das openCost-Github-Repository einsehbar ist und dynamisch erweitert wird (vgl. openCost-Github-Repository). Auf diese Weise findet es Verwendung als Referenz in der Arbeitspraxis der Publikationskostenverwaltenden. Der skizzierte Prozess der zentralen Publikationskostenerfassung verlangt eine über viele Bereiche hinweg funktionierende Verständigung – so auch für Institutspersonal mit wenig Publikationskenntnissen – bei der man sich auf das Glossar berufen kann. Dieses wird um Begriffe ergänzt, die auf Publikationsrechnungen auftauchen und synonym zu einer der in den Tabellen 5.2 und 5.3 genannten Bezeichnungen stehen. Die Synonyme werden mit Verweis auf das jeweilige standardisierte Kostentypen-Stichwort verzeichnet.

Die Festlegung der Kostentypen in den Tabellen 5.2 und 5.3 eröffnet einen weiteren konkreten praktischen Nutzen. Das standardisierte Vokabular, das in das openCost-Metadatenschema (ebd.) integriert ist, ermöglicht einen automatisierten Austausch

von Kosteninformationen und deren Vergleich. Ein Metadatenschema ist nicht nur in seiner Struktur, seinem Format, sondern auch in seinem Vokabular zu definieren. Um die Kosteninformation für den einrichtungsindividuellen Bedarf, aber auch interinstitutionell zur Verfügung zu stellen, werden diese als Metadateninformation der Publikation hinterlegt.

Die Abbildung 5.1 visualisiert, welche Informationen in den Metadaten einer Publikation enthalten sind. openCost stellt einen Metadatenbaustein eingebettet in die Gesamtinformationen zu einer Publikation dar, der inhaltlich ausschließlich kostenrelevante Daten trägt. Einzig der doi (Digital Object Identifier) als primärer Identifikator wird als nicht kostenspezifische Information im openCost-Datensatz wiederholt, um eine eindeutige Zuordnung des Publikationskostenbausteins zur jeweiligen Publikation sicherzustellen und so über openCost auch reine Kostendaten ohne weiteren Kontext austauschen zu können.

openCost hat in der ersten Projektphase zwei Metadatenschemata entworfen: Eines zur Beschreibung einzelner Artikel (opencost: publication) sowie ein weiteres zur Erfassung von (Transformations-)Verträgen und Mitgliedschaften (opencost: contract) (vgl. openCost-Github-Repository).[14]

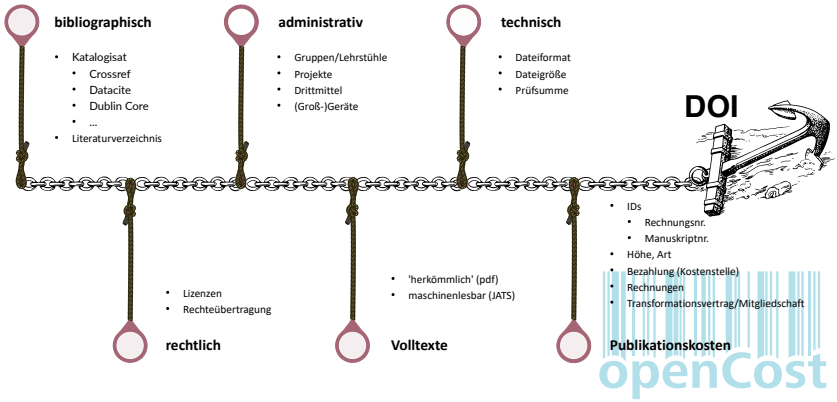


Abbildung 5.1: Metadatenbaustein openCost (Deinzer, Gernot & Wagner, Alexander), Lizenz: CC BY 4.0.

Die Metadaten für einzelne Artikel werden, z.B. in einem institutionellen Repositorium, der Hochschulbibliographie oder einem Forschungsinformationssystem, zum Artikeleintrag hinzugefügt. Zur Abbildung eines Vertrags erhält dieser einen eigenen Eintrag und wird durch persistente Identifikatoren eindeutig identifizierbar und mit dazugehörigen Einzelpublikationen verknüpft. Folglich existieren zwei zu differenzierende Informationsebenen.

Das Projekt hat zwar mit der Beschreibung von wissenschaftlichen Zeitschriftenartikeln und Open-Access-Verträgen begonnen, das Metadatenschema ist im strukturellen Aufbau aber hinreichend flexibel, um analog für andere Bereiche erweitert zu werden. Schlussendlich soll das Schema die Darstellung eines jeden Szenarios einer Finanzierung im Publikationsprozess ermöglichen.

Die in den Tabellen 5.2 und 5.3 unter „other" aufgeführten Kostenarten sind noch nicht in der aktuellen openCost-Version enthalten, werden aber im Projektfortgang geprüft und ergänzt. Es handelt sich um vorläufige Terme, deren Verwendungskontexte mit dem noch unvollständigen Kostenalphabet (vgl. Tab. 5.1) aufgezeigt werden, z.B. Erweiterungen des veröffentlichten Objekts durch „APC", „BPC", „CPC" usw. „APC" und Vergleichbare werden dann im openCost-Metadatenschema mit den Kostentypen „gold-oa" oder „hybrid-oa" gekennzeichnet. Auch die definierten Zusatzkosten (cover charge, permission, payment fee etc.) können für jede Variation des publizierten Objekts (BPC, CPC etc.) anfallen, die Schemabeschreibung bleibt konstant. Das Prinzip lässt sich folglich auf viele weitere Szenarien übertragen. Kostentypen, die im Schema insbesondere unter „other" separat genannt werden, können zudem mitunter schon in einer „Processing Charge" (vgl. „PC" in „APC" und Weiteren) enthalten sein. Liegt jedoch mit einer Verlagsrechnung eine andersartige, präzise Aufschlüsselung vor, in der z.B. die Kosten des wissenschaftlichen Lektorats explizit verzeichnet sind, können diese ebenfalls exakt dokumentiert werden. Eine Erweiterung des aktuellen openCost-Vokabulars wird hierfür angestrebt.

Auch das Vertragsschema ist hinlänglich flexibel zur Beschreibung unterschiedlicher Vereinbarungen: Transformationsverträge, Mitgliedschaften, Abonnements und weitere. Die Subskription von Zeitschriften stellt ebenfalls eine Vertragsform dar, obgleich sie nicht das Publizieren, sondern den Lesezugriff thematisiert. Die Ausgaben dafür sind jedoch Teil des Informationsbudgets einer Einrichtung. Obwohl openCost zunächst nur zur Darstellung von Publikationskosten konzipiert war, wird deutlich, dass die Flexibilität des Schemas nicht auf ausgewählte Individualszenarien beschränkt ist. Im Zuge der Subskription wird eine Gebühr für ein Abonnement berechnet, das den Lesezugriff für ein ausgewähltes Portfolio gewährt. Dieser Aspekt erfordert keine terminologische Erweiterung des openCost-Schemas, sondern kann durch den Kostentyp „read" abgebildet werden.

Gleichermaßen lässt sich das Vertragsschema zur Wiedergabe von Diamond-Open-Access-Szenarien verwenden. Das Vertragsobjekt enthält dann den Kostentyp „publish". Komplexe Modellvarianten sollen durch geeignete Erweiterungen ergänzt werden. Die in Tabelle 5.3 genannte „service fee" ist in diesem Zusammenhang gegebenenfalls noch zu allgemein gefasst, eine „editing charge" (Tabelle 5.2), um die redaktionellen Prozesse im eigenen Haus zu beschreiben, wird dann möglicherweise nicht auf Artikel-, sondern auf Vertragsebene nützlich.

Insgesamt können einige Kostentypen sowohl für Einzelpublikationen als auch für Vertragskonstrukte relevant werden: Dies betrifft nicht nur die zu leistende Mehrwertsteuer („vat"), sondern es wird sich auch im Einzelfall herauskristallisieren, ob eine Servicepauschale (vgl. „service fee" unter „other", Tabelle 5.3), wie sie z.B im Rahmen von DEAL zu leisten ist, artikel- oder vertragsbasiert abzurechnen ist. Der Service-Aspekt kann des Weiteren in anderen Fällen schon im „Processing"-Part einer „Processing Charge" enthalten sein.

Mit dem Kostentyp „infrastructure charge" soll ebenfalls ein weiteres Szenario exemplarisch aufgezeigt werden, das über die Finanzierung einer Publikation oder eines dazugehörigen Vertrags hinausgeht, technisch jedoch mit dem openCost-Vertragsschema

abgebildet werden kann. Es geht dabei um Kostenbeiträge durch eine wissenschaftliche Einrichtung zur Finanzierung einer wissenschaftsgeleiteten Open-Science-Infrastruktur, z.B. die gemeinschaftliche Finanzierung des *DOAJ* oder die kooperative Förderung der *EZB*.

Weitere Diamond-Finanzströme und alle Aspekte, um einen Verlag betriebswirtschaftlich zu finanzieren, können aufgeschlüsselt werden. Dies ist wichtig, um davon abgeleitet Preise für Einzelpublikationen sowie für Publish & Read-Anteile zu errechnen.

Die Initiativen *PLOS* (PLOS 2024), *arXiv* (arXiv 2023: 10) oder die *MPDLS* (MPDL Services gGmbH 2023: 2) mit den Betriebskosten zur Servicepauschale der *DEAL*-Verträge liefern entsprechend Beispiele der Dokumentation budgetärer Verteilungen für den Publikationsprozess. So untergliedert beispielsweise der Open-Access-Verlag *PLOS* seine Gesamtausgaben in die drei Bereiche „Publishing", „General and Administrative" sowie „Fee Assistance" (PLOS 2024). „Publishing" setzt sich zusammen aus Redaktion, Produktion und Technologie: „Efforts related to editorial and production, operations and advocacy, including staffing, activity tools" (ebd.). Der Bereich „General and Administrative" inkludiert nichtpublikationsbezogene Kosten für Personal und ausgelagerte Dienstleistungen in den Bereichen Personalwesen, Recht und Buchhaltung, zudem umfasst er auch Mieten und Bankgebühren. „Fee Assistance" beschreibt die Gebührenhilfe für Autor:innen, die nicht in der Lage sind, die Kosten für die Veröffentlichung ihrer Artikel ganz oder teilweise zu zahlen.

Die Forderung nach Transparenz geht einher mit einem Blick ins Detail und einer darzustellenden Realität, die komplex ist. Dabei ist jedoch darauf zu achten, dass die Darstellung der Kosten im Publikationsprozess nicht beliebig oder unüberschaubar wird. Eine zu ausgedehnte Beschreibung potenziell relevanter Kosten, die nur im weitesten Sinne einen Publikationsbezug haben, führt zu einem Verlust des Fokus'. Der Wunsch nach Detailtreue ist zudem gegen die Fehleranfälligkeit einer zu kleinteiligen Darstellung und den damit einhergehenden Erfassungsaufwand

abzuwägen. In einer Datenflut geht das eigentliche Ziel – die Vergleichbarkeit als Vorteil eines Standards und Grundlage für auswertbare Ergebnisse – verloren.

2.3 Systematisierung, Standardisierung & Automatisierung

Die Open-Access-Transformation wird fortan Einfluss auf die Ausgestaltung der Geschäftsmodelle nehmen und weitere Bedingungen und Szenarien hervorbringen. Das von openCost entwickelte Metadatenschema ermöglicht eine strukturierte Abbildung, Erfassung und Abfrage von Publikationskosten, die in diesem Zusammenhang anfallen.

Das Metadatenformat wurde in Kooperation mit nationalen und internationalen Partnern erarbeitet, um den langfristigen, variierenden Anforderungen der Kostenauswertung gerecht zu werden. Ziel ist die Bereitstellung einer zuverlässigen und nachhaltigen Datenbasis, die nicht spezifisch für individuelle Anfragen konzipiert ist, sondern allgemein einsetzbar bleibt.

So kann eine Einrichtung Publikationskostendaten standardisiert erfassen und sich bei der internen Aufbereitung von Informationen an den Inhalten und Begriffen des openCost-Metadatenschemas orientieren. Die Vorteile eines solchen Standards sind jedoch nicht nur für die individuelle Nutzung relevant: Ein expliziter Fokus von openCost liegt auf der offenen Bereitstellung und dem automatisierten Austausch der Kostendaten. Es existieren verschiedene Anwendungsszenarien, in denen dieser Austausch notwendig ist, und die Abläufe können durch den openCost-Standard vereinfacht werden. Sowohl innerhalb einer Einrichtung als auch zwischen Institutionen sind so Kostenberichte sowie Vergleiche möglich. Die im openCost-Format gespeicherten Kostendaten einer Einrichtung können durch Drittinstanzen abgerufen werden, z.B. zur Verfügbarmachung von Publikationskosten auf Informationsplattformen oder zur Berichterstattung für Forschungsfördernde.

Ein konkretes Anwendungsszenario in der deutschen Wissenschaftslandschaft ist das bereits genannte DFG-Projekt „Open-Access-Publikationskosten", das finanzielle Unterstützung für Open-Access-Publikationen in wissenschaftlichen Einrichtungen bereitstellt. Hierfür müssen jährlich Nachweise an die DFG erbracht werden, die den Mitteleinsatz für Open-Access-Publikationen belegen.[15]

Durch die Nutzung einer etablierten technischen Schnittstelle können die openCost-Daten kontinuierlich abgerufen werden, sodass alle am Publikationsprozess Beteiligten von einer aktuellen Datenlage profitieren und etwa Preise vergleichen oder Ausgaben prüfen können. Regelmäßige Updates ermöglichen eine umfassendere Übersicht und Planbarkeit. Aktualisierungen, Korrekturen oder Ergänzungen lassen sich automatisiert und zeitnah umsetzen – ein wesentlicher Beitrag zur Effizienzsteigerung in der Forschungsdokumentation. Jener kontinuierliche Datenfluss könnte perspektivisch auch den jährlichen Nachweisprozess zum beschriebenen DFG-Programm „Open-Access-Publikationskosten" ersetzen.

Für alle potentiellen Anwendungsfälle bedeutet dies, dass die manuelle Erstellung spezifischer Listen durch einen automatisierten und regelmäßigen Datenfluss ersetzt wird, wobei die Aggregation der Publikationskostendaten direkt aus dem erfassenden System erfolgt. Repositorien bieten sich hierfür beispielsweise an, da sie bereits eine Vielzahl von hochwertigen Metadaten in verschiedenen Formaten bündeln, die die Publikation an ihrem Speicherort beschreiben. Außerdem archivieren Repositorien die veröffentlichten Werke und sorgen dafür, dass sie weltweit dauerhaft verfügbar sind. Dies ist der Ort, an dem das umfassende Wissen über eine Publikation zusammenkommt.

Durch diese Automatisierung und Standardisierung lassen sich signifikante Einsparungen in den Arbeitsabläufen erzielen (vgl. auch Stein 2024: 466 ff.).

openCost liefert somit ein nachhaltiges und maschinenlesbares Format zur langfristigen Bereitstellung und effizienten Nutzung, das den *FAIR*[16]-Kriterien entspricht (Wilkinson 2016: 4).

3. Schlussbetrachtung

Die Auseinandersetzung mit Publikationskosten zeigt, wie komplex und vielschichtig die Finanzierung wissenschaftlicher Veröffentlichungen ist. Damit sich die für die wissenschaftliche Kommunikation so bedeutende Open-Access-Transformation auch finanziell nachhaltig gestaltet, sind klare Transparenz und nachvollziehbare Kostenstrukturen von zentraler Relevanz. Es bedarf systematischer Anpassungen, um den Wandel zu Open Access zu ermöglichen und gleichzeitig die Transparenz der Kosten für die Veröffentlichung von Forschungsergebnissen zu verbessern.

Das Projekt openCost stellt hierbei einen bedeutenden Schritt dar, um die vielschichtigen Kostentypen und -modelle standardisiert abzubilden und automatisiert zugänglich zu machen. Durch jene Standardisierung und Automatisierung von Publikationskosten wird ein Beitrag zu einer nachhaltigen und gerechten Wissenschaftskultur geleistet, die sowohl den Zugang zu Wissen fördert als auch eine faire Finanzierung gewährleistet.

Wissenschaftliche Bibliotheken spielen eine Schlüsselrolle in diesem Transformationsprozess, indem sie als zentrale Instanzen die Verantwortung für die Rechnungsbearbeitung und das Publikationsmanagement übernehmen. Durch die Implementierung eines einheitlichen Metadatenschemas und die Nutzung von definierten Schnittstellen wird es möglich, Publikationskosten systematisch zu erfassen und auszutauschen. Diese Maßnahmen fördern nicht nur die Effizienz im Publikationsprozess, sondern unterstützen auch die wissenschaftliche Gemeinschaft dabei, fundierte Entscheidungen in der Publikationswahl zu treffen und Budgetierungsprozesse vorausschauend zu gestalten.

Die Flexibilität des Metadatenschemas, das an zukünftige Anforderungen angepasst werden kann, stellt sicher, dass es auch in einem sich dynamisch verändernden Umfeld relevant bleibt. Das Publikationskostenalphabet kann erweitert werden. Durch die Implementierung automatisierter Prozesse wird der manuelle Aufwand reduziert und eine kontinuierliche Aktualisierung der Kostendaten gewährleistet. Dies verbessert nicht nur die

Datenqualität, sondern unterstützt auch die langfristige Nachvollziehbarkeit der Ausgaben.

Der Fokus auf ein offenes und interoperables System ist ein wichtiger Schritt in Richtung einer nachhaltigeren und transparenteren Forschungslandschaft.

Danksagung

Diese Publikation wurde gefördert durch die Deutsche Forschungsgemeinschaft (457354095) im Rahmen des Projekts „openCost".

Anmerkungen

[1] Vgl. den Beitrag „Datentracking – Gefahr für die Wissenschaft?" von Bernhard Mittermaier in diesem Band.
[2] Zum Projekt vgl. https://www.opencost.de/, eingesehen am 17. November 2024.
[3] Welche Parameter das sein können, beschreibt Bernhard Mittermaier (2022: 6ff.) in einem Schalenmodell.
[4] „Elektronische Zeitschriftenbibliothek" im Folgenden „EZB".
[5] „Directory of Open Access Journals" im Folgenden „DOAJ".
[6] „Deutsche Forschungsgemeinschaft" im Folgenden „DFG".
[7] Vgl. Kap 2.1.1 zur Erläuterung.
[8] DFG-Projektnummer: 457354095.
[9] „Open Archives Initiative Protocol for Metadata Harvesting" im Folgenden „OAI-PMH".
[10] *Article Processing Charge* im Folgenden „APC".
[11] EPC und IPC sind in der Tabelle 5.1 mit dem Zusatz „(2)" markiert, da sie in der Darstellung durch openCost anders verarbeitet werden als die mit einer „(1)" versehenen publizierten Objekte (vgl. Kapitel 2.2).
[12] Tabelle 5.2 verzeichnet Kostentypen, die mit (1) markierte publizierte Objekte in Tabelle 5.1 betreffen.
[13] Tabelle 5.3 beinhaltet die Kostentypen, die gemäß der in Tabelle 5.1 mit (2) markierten kostentragenden Entitäten Kostenkontexte beschreiben.
[14] Die in Tabelle 5.1 mit (1) markierten Gebührenbezeichnungen zur Open-Access-Freigabe eines spezifischen publizierten Objekts finden sich in der Darstellung des openCost-Metadatenschemas für

einzelne Artikel wieder. Das im Metadatenschema für Einzelartikel definierte Vokabular ist in der Tabelle 5.2 aufgeführt. Analog liegen die in Tabelle 5.1 mit (2) gekennzeichneten Entitäten dem open-Cost-Metadatenschema für Verträge zugrunde. Die Liste zulässiger Kostentypen enthält Tabelle 5.3.

[15] Das Forschungszentrum Jülich ist mit dem jährlichen Monitoring zum Förderprogramm „Open-Access-Publikationskosten" durch die Deutsche Forschungsgemeinschaft (DFG) beauftragt. Dieses dokumentiert die Open-Access-Publikationen sowie verausgabten Publikationsmittel einer wissenschaftlichen Einrichtung und soll zur Optimierung der dortigen Datenlage beitragen. Es dient der DFG als Bericht eines Förderjahres, in dem Open-Access-Publikationen – in der zweiten Förderphase ausschließlich jene, die aus DFG-Förderungen hervorgehen – bezuschusst werden. Das Datenschema zum Monitoring enthält Publikationsmetadaten, inklusive der Finanzdaten.

[16] Das Akronym „FAIR" steht für die Kriterien „findable", accessible", „interoperable" und „reusable". Diese dienen als Leitprinzipien zur Organisation von Forschungs(meta)daten.

Literatur

Akademischer Senat der Humboldt-Universität zu Berlin (2014) *Grundsätze zum Umgang mit Forschungsdaten an der Humboldt-Universität zu Berlin (Forschungsdaten-Policy) vom 08.07.2014*. https://www.cms.hu-berlin.de/de/dl/dataman/hu-fdt-policy/, Zugriff am 21.10.2024.

arXiv (2024) https://arxiv.org/, Zugriff am 25.08.2024.

arXiv (2023) Annual Report 2023. https://info.arxiv.org/about/reports/2023_arXiv_annual_report.pdf, Zugriff am 28.10.2024.

Budapest Open Access Initiative (2002) https://www.budapestopenaccessinitiative.org/read/, Zugriff am 23.08.2024.

DEAL Open Access Services (DEAS) gGmbH (2024) DEAL Konsortium. https://deal-konsortium.de, Zugriff am 28.08.2024.

Deinzer, Gernot und Cornelia Lang (2024) „All in One: Informationsbudget an der Universität Regensburg. Umsetzung, Strukturen und Geschäftsprozesse". *ZfBB Zeitschrift für Bibliothekswesen und Bibliographie* 71(2): 91–101. https://doi.org/10.3196/186429502471249.

Deutsche Forschungsgemeinschaft (DFG) (2024) Neue Dynamik bei Diamond Open Access. Eine Ausschreibung im Rahmen des LIS-

Förderprogramms „Infrastrukturen für wissenschaftliches Publizieren". https://www.dfg.de/resource/blob/324442/db52a2a58f84c92b1c3d7cf9cb2eaf41/call-diamond-open-access-de-data.pdf, Zugriff am 28.10.2024.

DOAJ, Directory of Open Access Journals: https://doaj.org/, Zugriff am 21.08.2024.

Drößler, Stefan (2022) „Open Access. Technikgetriebene Gesellschaftsutopie für die Transformation des wissenschaftlichen Publikationssystems". In Nicola Mößner/Klaus Erlach (Hrsg.): *Kalibrierung der Wissenschaft. Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis*, Bielefeld: Transcript, S. 79–101. https://doi.org/10.14361/9783839462102-004.

EZB, Elektronische Zeitschriftenbibliothek: https://ezb.uni-regensburg.de/, Zugriff am 21.08.2024.

Forschungszentrum Jülich (2024) Monitoring des DFG-Förderprogramms „Open-Access-Publikationskosten". https://www.fz-juelich.de/de/zb/open-science/open-access/monitoring-dfg-oa-publikationskosten, Zugriff am 30.10.2024.

Frick, Claudia und Christian Kaier (2020) „Publikationskosten für Zeitschriftenartikel abseits von Open-Access-Publikationsfonds – Lost in Transformation?" *O-Bib. Das Offene Bibliotheksjournal. Herausgeber VDB*, 7(2): 1–15. https://doi.org/10.5282/o-bib/5586.

GEPRIS. Geförderte Projekte der DFG: *openCost: automatisierte, standardisierte Lieferung und offene Bereitstellung von Publikationskosten und Verlagsvereinbarungen*: https://gepris.dfg.de/gepris/projekt/457354095, Zugriff am 26.08.2024.

HMC-Webseite: Helmholtz Metadata Collaboration. https://helmholtz-metadaten.de/de, Zugriff am 21.10.2024.

Mittermaier, Bernhard (2022) „Das Informationsbudget: Konzept und Werkstattbericht". *O-Bib. Das Offene Bibliotheksjournal Herausgeber VDB*, 9(4): 1–17. https://doi.org/10.5282/o-bib/5864.

MPDL Services gGmbH (2023) Informationen zur Servicepauschale. Hintergründe, Verwendung und Perspektiven. https://keeper.mpdl.mpg.de/f/9b69151e07444e6aa1b0/?dl=1, Zugriff am 28.10.2024.

NFDI-Webseite (2024) Nationale Forschungsdateninfrastruktur. https://www.nfdi.de/, Zugriff am 21.10.2024.

OAI-PMH: Open Archives Initiative Protocol for Metadata Harvesting: https://www.openarchives.org/pmh/, Zugriff am 26.08.2024.

open-access.network: Glossar. https://open-access.network/informieren/glossar#c6188, Zugriff am 21.10.2024.

Open-Access-Publikationskosten: Förderprogramm der Deutschen Forschungsgemeinschaft: https://www.dfg.de/de/foerderung/foerdermoeglichkeiten/programme/infrastruktur/lis/lis-foerderangebote/open-access-publikationskosten, Zugriff am 23.08.2024.

OpenAlex: https://openalex.org/, Zugriff am 20.06.2025.

openAPC-Webseite: https://openapc.net/, Zugriff am 26.08.2024.

openCost-Github-Repository: https://github.com/opencost-de/opencost/tree/main/doc, Zugriff am 26.08.2024.

openCost-Webseite: https://www.opencost.de/projekt/, Zugriff am 26.08.2024.

Open policy finder: https://openpolicyfinder.jisc.ac.uk/, Zugriff am 20.06.2024.

Open Science Festival (2024) https://converia.uni-mainz.de/frontend/index.php?folder_id=924&page_id=, Zugriff am 21.10.2024.

Pampel, Heinz (2019) *Auf dem Weg zum Informationsbudget. Zur Notwendigkeit von Monitoringverfahren für wissenschaftliche Publikationen und deren Kosten*. Arbeitspapier, Potsdam. https://doi.org/10.2312/os.helmholtz.006.

Pampel, Heinz; Bargheer, Margo; Bertelmann, Roland; Kaden, Ben; Kindling, Maxi; Schobert, Dagmar; Wrzesinski, Marcel (2024) „Thesen zur Zukunft des wissenschaftsgeleiteten Open-Access-Publizierens – Entwicklung und Diskussion in einem partizipativen Prozess". *Bibliothek Forschung und Praxis*, 48(2): 322–333. https://doi.org/10.1515/bfp-2024-0027.

PLOS (2024) 2022 Financial Overview. https://plos.org/financial-overview/, Zugriff am 21.10.2024.

SCOSS (2024) What is Open Infrastructure? https://scoss.org/what-is-scoss/defining-open-infrastructure/, Zugriff am 28.10.2024.

Springer Support (2024) Article Processing Charges (APC). https://support.springer.com/en/support/solutions/articles/6000211135-article-processing-charges-apc-, Zugriff am 21.10.2024.

Springer (2023) Open Access. Article Processing Charges. https://www.springer.com/gp/authors-editors/authorandreviewertutorials/open-access/article-processing-charges/10286526, Zugriff am 30.10.2024.

Stein, Lisa-Marie (2024) „Kommerzielles Wissen? Transparent publizieren mit openCost". *b.i.t. online*, 27(5): 461–469. https://doi.org/10.3204/PUBDB-2024-04986.

Wagner, Alexander (2024) „How much is the fish? openCost for cost transparency and sustainable and FAIR scientific publishing". 3. Open Science Festival, Mainz, Germany, 17. – 18.09.2024. https://doi.org/10.3204/PUBDB-2024-05830.

Wagner, Alexander; Stein, Lisa-Marie; Bartlewski, Julia; Broschinski, Christoph; Deinzer, Gernot; Pieper, Dirk; Schweighofer, Bianca; Sippl, Colin; Weisheit, Silke (2023) *Publication cost transparency within the Helmholtz Community and beyond: openCost metadata for different publication types – from APCs up to research data. HMC Conference: Better Research Through Better Metadata, 10. – 12.10.2023*. https://doi.org/10.3204/PUBDB-2023-06134.

Wiley Authorservices (2024) Article Publishing Charges. https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/article-publication-charges/index.html, Zugriff am 30.10.2024.

Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, IJsbrand Jan u.a. (2016) „The FAIR Guiding Principles for scientific data management and stewardship". *Sci Data* 3, 160018. https://doi.org/10.1038/sdata.2016.18.

Wissenschaftsrat (2022) *Empfehlungen zur Transformation des wissenschaftlichen Publizierens zu Open Access*, Köln. https://doi.org/10.57674/fyrc-vb61.

TEIL II

Die Frage nach der Expertise

TEIL II

Die Frage nach der Expertise

Klaus Erlach und Nicola Mößner

Die Bewertung wissenschaftlicher Expertise ist keineswegs trivial und nichtsdestotrotz genauso wichtig wie die wissenschaftliche Tätigkeit selbst. Erst durch Anerkennung in der jeweiligen wissenschaftlichen Community wird ein wissenschaftliches Ergebnis eben genau dazu, andernfalls geht es verloren. Man muss nicht Relativist sein, um der Aussage zuzustimmen, dass Wissenschaft nur dasjenige ist, was Wissenschaftler:innen dafür halten.

Natürlich gibt es zahlreiche allgemeine Kriterien für die Bestimmung dessen, was wissenschaftliches Arbeiten bedeutet. Bekannt ist in diesem Zusammenhang insbesondere die Frage nach dem sogenannten Demarkationskriterium, welches also Wissenschaft von bloßer Pseudowissenschaft unterscheiden soll (vgl. z.B. Gordin 2021; Pigliucci & Boudry 2013; Chalmers 2007). Aus theoretischer Perspektive haben Generationen von Wissenschaftstheoretiker:innen vor allem jedoch gezeigt, dass wissenschaftliches Wissen prekär ist. Wissenschaftliche Ergebnisse sind also trotz (oder je

Zitationsvorschlag für dieses Kapitel:
Erlach, K. und Mößner, N. 2026. Teil II: Die Frage nach der Expertise. In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* S. 131–137. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.g.
License: CC BY 4.0

nach Sichtweise auch gerade wegen) all der Methodik fallibel und nicht absolut und dauerhaft gültig. Und aus praktischer Perspektive haben zahlreiche Krisen wissenschaftlicher Verlässlichkeit – angefangen mit explizitem Fehlverhalten wie Betrug (Datenfälschung) und Ergebnisverfälschung (p-Hacking) bis hin zu impliziten, dem Wissenschaftssystem immanenten Fehlleistungen wie der Replikationskrise (vgl. Übersicht in Erlach 2022) – diese Problematik bestätigt.

Aus diesen hier nur skizzenhaft angedeuteten Gründen darf man schließen, dass das Vorgehen bei der Wissenschaftsevaluation durchaus äußerst relevante Rückwirkungen auf die Erzeugung wissenschaftlichen Wissens haben kann. Solange sich Evaluierungsverfahren nicht ändern, mag das unproblematisch erscheinen. Nun ist aber seit geraumer Zeit ein Skaleneffekt hinzugekommen, der auch Einfluss auf die Bewertungsverfahren genommen hat. Die Expansion des Wissenschaftssystems im 20. Jahrhundert hat nicht nur dazu geführt, dass es mehr lebende als im Verlauf der Menschheitsgeschichte verstorbene Wissenschaftler:innen gibt (oder gegeben hat, je nach Fachgebiet; vgl. Solla Price 1974), sondern auch dazu, dass die Menge der dokumentierten Forschungsergebnisse entsprechend explodiert ist.[1] Deshalb scheint die Steigerung der Effizienz der Bewertungsverfahren – und damit ein Trend hin zu quantitativen Metriken – geradezu zwingend. So ist es dazu gekommen, dass die in den Bibliothekswissenschaften entwickelte Bibliometrie mittlerweile eine wichtige Basis der Beurteilung wissenschaftlicher Leistungen (ausgedrückt als h-Index, Impact Faktor usw.) bildet (vgl. Retzlaff 2022).

Grundlage für die bibliometrisch gestützte Evaluation bilden in der Regel kommerzielle Datenbanken wie Scopus und Web of Science. Dabei zeigt sich ein wesentlicher Grund für das Festhalten der wissenschaftlichen Gemeinschaft an den problematischen Praktiken des elektronischen Publizierens: das Interesse und der vermeintliche Bedarf an einfachen, weil quantitativen Evaluationsmaßstäben. Metrische Analysen setzen, wie angemerkt, entsprechend gut gefüllte Datenbanken voraus. Diese sind zum derzeitigen Stand der Dinge aber fest in kommerzieller Hand.

Der riesige Bestand wissenschaftlicher Publikationen hat letztlich dazu geführt, dass die Wissenschaftsevaluation abhängig von Informationsdienstleistern geworden ist, deren Hauptinteresse mit dem übereinstimmt, was auch die Metriken messen: das schiere Mengenwachstum an Publikationen. Die „Waage der Beurteilung von Expertise" neigt sich den quantitativen Metriken und dem Gewicht zahlloser, nicht weiter differenzierter Einzelergebnisse zu (vgl. Abb., linke Seite).

Indem quantitative Metriken jegliche Inhalte völlig ausblenden, wird wissenschaftliches Wissen zum bloßen Datum. Dann ist es aber von nichtwissenschaftlichem Wissen bis hin zu Verschwörungstheorien formal nicht mehr unterscheidbar, weil jegliche Kriterien der Wissenschaftlichkeit in der metrischen Bewertung ausgeblendet sind. Sie werden allenfalls dem vorgängigen Review-Prozess unterstellt, aber nicht mehr eigens thematisiert. Zugleich müssen sich die individuellen Wissenschaftler:innen diesem System der Metrisierung unterwerfen, da

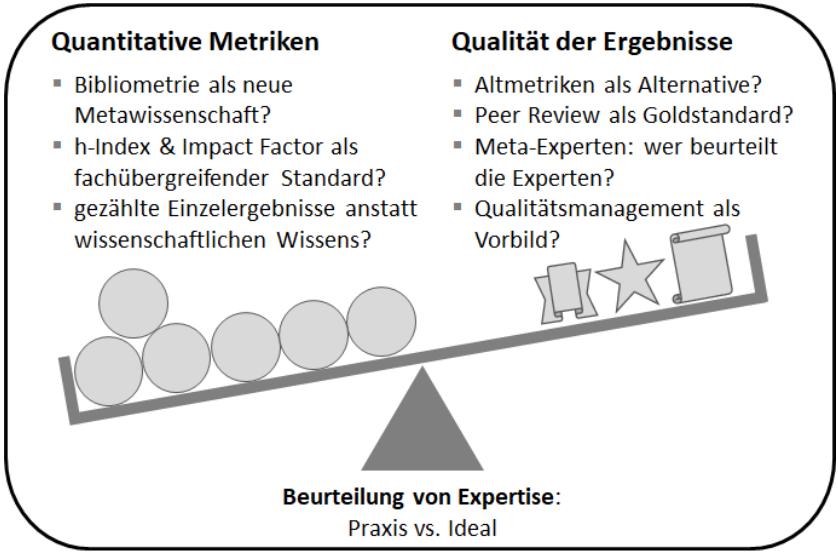


Abbildung: Die Beurteilung wissenschaftlicher Expertise in der Praxis bevorzugt quantitative Metriken.

es von hoher Relevanz für die Etablierung im Beruf und ihre wissenschaftliche Karriere ist. Die Maxime „Publish or Perish" bringt das sehr treffend zum Ausdruck. Die dadurch geförderte Leichtgewichtigkeit der eigentlichen Ergebnisqualität (vgl. Abb., rechte Seite) bräuchte also neue Impulse, um in der Wissenschaftsevaluation weiterhin eine Rolle spielen zu können.

Dieser Weg von der Dominanz quantitativer Verfahren hin zur Beleuchtung der Potentiale qualitativer Verfahren soll in diesem Teil II *Die Frage nach der Expertise* in vier Beiträgen anhand folgender Fragen näher beleuchtet werden:

1. Welche grundsätzlichen Möglichkeiten zur Evaluierung wissenschaftlicher Expertise sind jeweils unter welchen Bedingungen zur Anwendung geeignet?
2. Wie funktioniert die wissenschaftsinterne Selbstkorrektur auf Basis von Evaluierungsergebnissen?
3. Wie kann aus wissenschaftsexterner Perspektive beurteilt werden, welchen Expert:innen und welcher Expertise vertraut werden kann?
4. Wie kann die Wissenschaftsevaluation mit Blick auf das Qualitätsmanagement in der Industrie weiterentwickelt werden?

Dieser Blick auf die Möglichkeiten, Grenzen und optionalen Zukunftsszenarien der Wissenschaftsevaluation unter den Bedingungen der Digitalisierung legt die Basis für die in Teil III erarbeiteten Lösungsvorschläge zu Autonomie und Offenheit der Wissenschaften.

Im ersten Beitrag fragt Bernhard Mittermaier, ob es einen Königsweg bei den Evaluierungsmethoden gibt. Zunächst stellt er die klassische Methode des Peer Review, die etablierten bibliometrischen Analysen und eher neueren, Altmetrics genannten Methoden vor. In einer Auswertung vorhandener Studien werden dann die jeweiligen Problemfelder sowie die Eignung der drei Verfahren für verschiedene Anwendungsfelder diskutiert. Im Ergebnis zeigt sich das *Peer Review* als Mittel der Wahl bei kleinskaligen Anwendungen (Manuskripte, einzelne Personen). Bei der Evaluierung von Forschergruppen ist die Unterstützung durch *Bibliometrie* sinnvoll; bei nationalen Auswertungen ist dies unabdingbar, zumal

dabei auch die Zielsetzungen andere sind. *Altmetrics* eignen sich den Studien zufolge nicht zur Forschungsevaluierung. Ihr Vorteil liegt aber darin, dass eine gewisse Vorhersagekraft für den späteren Zitationserfolg von Publikationen besteht.

Im zweiten Beitrag befassen sich Andreas Kaminski und Jörn Wiengarn mit dem Dilemma, dass bei fachlichem Dissens für die Auswahl der ‚richtigen' Expert:innen eine Expertise erforderlich ist, welche die widerstreitenden Fachergebnisse vergleichend bewerten können muss. Also gerade dann, wenn sich Expert:innen in praktisch relevanten Fragen nicht einig sind, stellt sich für Laien die Frage, wer tatsächlich als Experte gelten kann und wie groß dessen Expertise einzuschätzen ist. Wie aber können Laien in solchen Situationen die Expertise anderer einschätzen, ohne doch selbst Expert:innen zu sein? Für die Autoren läuft dies letztlich auf die Frage hinaus, welchen Expert:innen aus rationalen Gründen vertraut werden kann. Bisher in der Philosophie erarbeitete Vorschläge über das das Verhältnis von Vertrauen und Wissen scheitern daran, dass die Fundierung von Vertrauen durch vertrauensunabhängiges Wissen nicht durchführbar ist. Daher stellen die Autoren ein alternatives Modell vor, nach dem Vertrauen und Wissen in einem pervasiven Zusammenhang zueinander stehen. Dieser zeigt sich doppelt in der Vertrauensgeladenheit von Wissen einerseits und dem im Wissen eingekapselten Vertrauen andererseits. Anhand des Fallbeispiels der Debatte über den Ursprung von SARS-CoV-2 zeichnen sie die Schwierigkeiten des Fundierungsmodells nach und illustrieren die Plausibilität des alternativ vorgeschlagenen Pervasivitätsmodells.

Im dritten Beitrag geht Klaus Erlach auf das Wunder objektiver Wissenschaftler-Beobachtung ein. Die nähere Betrachtung von sechs grundlegenden Bewertungsverfahren zeigt, dass quantitative Verfahren für ihre „objektive" Messung auch einen Maßstab mit normativer Sollwert-Vorgabe benötigen, während qualitative Verfahren intersubjektiv validiert werden können. Ein Blick auf die historische Entwicklung der Methoden des Qualitätswesens in der Industrie zeigt, dass hier eine zum Wissenschaftsbetrieb gegenläufige Entwicklung von rein quantitativen

Qualitätsprüfungsmethoden hin zu qualitativ arbeitenden Qualitätsmanagementsystemen stattgefunden hat. Abschließend wird aufgezeigt, inwiefern die Wissenschaftsevaluation in der Philosophie sich davon anregen lassen kann.

Im letzten Beitrag geht Nora Hangel mit einem empirischen Ansatz der Frage nach, wie die Selbstkorrektur innerhalb der Wissenschaft funktioniert. Dabei hat sie in Einzelinterviews und in Diskussionsgruppen Wissenschaftsphilosoph:innen und experimentell arbeitende Sozialwissenschafter:innen gebeten, über ihre epistemischen Kriterien zu sprechen, wenn sie bei Peer-Review-Prozessen oder Selbstkorrektur wertende Entscheidungen treffen. Von der Klarheit der Formulierungen bis zur Konsistenz der Argumentation über verschiedene Karrierestufen hinweg zeigte sich dabei eine breite Überschneidung der epistemischen Kriterien. Im Unterschied zu den Philosoph:innen nutzen die empirischen Wissenschaftler:innen zusätzlich eine Methoden-Checkliste. Sie zeigt die Kombination aus allgemeingültigen Kriterien der Wissenschaftlichkeit einerseits und mehr oder weniger strikten fachspezifischen Kriterien der Methodologie andererseits. Was die Selbstkorrektur auf der Ebene der Wissenschaftssystems anbelangt, so scheinen die nach der Replikationskrise eingeführten Praktiken der Open Science in den betrachteten Sozialwissenschaften eine positive Wirkung gezeigt zu haben. Der Artikel endet mit einem Ausblick auf weitere Möglichkeiten der Selbstkorrektur, insbesondere in der Philosophie.

Anmerkung

[1] Ob das im gleichen Maße für das valide wissenschaftliche Wissen gilt, hängt stark vom Wissensbegriff ab.

Literatur

Chalmers, Alan F. (2007) Wege der Wissenschaft – Einführung in die Wissenschaftstheorie. 6. Aufl., Springer. https://doi.org/10.1007/978-3-540-49491-1.

Erlach, Klaus (2022) „Die Industrialisierung des Wissenschaftsbetriebs. Effiziente Erzeugung, Bereitstellung und Bewertung von Wissen", in: Mößner, N.; Erlach, K. (Hg.), Kalibrierung der Wissenschaft. transcript, S. 183–206. https://doi.org/10.14361/9783839462102-008.

Gordin, Michael D. (2021) On the Fringe: Where Science Meets Pseudoscience. Oxford University Press.

Pigliucci, Massimo; Maarten, Boudry (Hg.) (2013) Philosophy of Pseudoscience – Reconsidering the Demarcation Problem. University of Chicago Press.

Retzlaff, Eric (2022) „Wer bewertet mit welchen Interessen wissenschaftliche Publikationen? Eine Skizzierung des Einflusses kommerzieller Interessen auf die Forschungsoutput-Bewertung", in: Mößner, N.; Erlach, K. (Hg.), Kalibrierung der Wissenschaft. transcript, S. 139–162. https://doi.org/10.14361/9783839462102-006.

Solla Price, Derek J. de (1974 [engl. 1963]) Little Science, Big Science. Von der Studierstube zur Großforschung. Suhrkamp. https://doi.org/10.7312/pric91844.

KAPITEL 6

Peer-Review, Bibliometrie, Altmetrics – gibt es einen Königsweg bei Evaluierungen?[1]

Bernhard Mittermaier

1. Einleitung

Zur Evaluierung wissenschaftlicher Leistung lassen sich grundsätzlich drei Ansätze unterscheiden. Ein *Peer-Review* (Reinhart & Schendzielorz 2024) ist erstens eine Bewertung eines Manuskripts oder eines Forschungsvorschlags. Dieser wird von einer oder mehreren Personen durchgeführt, die über ähnliche Fachkenntnisse wie die Autoren der Arbeit verfügen, d.h. von Gleichrangigen (Weller 2001). Der Begriff umfasst auch die retrospektive Bewertung der vergangenen Leistung eines Wissenschaftlers oder einer Gruppe von Wissenschaftlern (Derrick 2019).[2] *Bibliometrie* (Ball 2022) ist zweitens ein wissenschaftliches Verfahren zur quantitativen Analyse von Publikationen und anderen

Zitationsvorschlag für dieses Kapitel:
Mittermaier, B. 2026. Peer-Review, Bibliometrie, Altmetrics – gibt es einen Königsweg bei Evaluierungen? In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* S. 139–167. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.h. License: CC BY 4.0

Kommunikationsformen in der Wissenschaft. Sie nutzt statistische Methoden, um Muster und Trends in der wissenschaftlichen Literatur zu identifizieren, wie die Anzahl der Veröffentlichungen, Zitationshäufigkeiten und Netzwerke von Autoren und Institutionen. Ziel ist es, die Entwicklung von Forschungsfeldern zu verstehen, wissenschaftliche Leistung zu bewerten und Entscheidungsgrundlagen für Forschungsförderung und -politik zu schaffen. (Haustein & Larivière 2015) *Altmetrics* sind drittens alternative Metriken zur Messung der wissenschaftlichen Wirkung und Rezeption von Forschungsarbeiten (Tunger & Meier 2018). Im Gegensatz zu traditionellen bibliometrischen Analysen erfassen Altmetrics eine Vielzahl von Online-Aktivitäten wie Erwähnungen in sozialen Medien, Downloads, Leserzahlen, Kommentare und Verlinkungen in Blogs und Nachrichtenartikeln. Altmetrics ergänzen traditionelle bibliometrische Methoden, indem sie die Aufmerksamkeit und das Engagement der breiten Öffentlichkeit und der wissenschaftlichen Gemeinschaft für Forschungsarbeiten messen und somit ein breiteres Verständnis der wissenschaftlichen Wirkung ermöglichen (Tunger & Meier 2022).

Dieser Beitrag beleuchtet zunächst die drei vorstehend genannten Verfahren etwas detaillierter und diskutiert Kritikpunkte an den jeweiligen Verfahren. Anschließend wird eine Auswahl von Studien vorgestellt, die Bibliometrie und Altmetrics sowie Peer-Review und bibliometrische Indikatoren (seltener: altmetrische Indikatoren) miteinander vergleichen. Den Abschluss bildet eine Diskussion der Ergebnisse.

2. Peer-Review

Die wahrscheinlich wichtigste Art des Peer-Reviews findet statt, wenn ein Manuskript bei einer Zeitschrift eingereicht wird (Paltridge 2017). Zeitschriften, die ein Peer-Review-Verfahren eingeführt haben, werden als „peer reviewed journals“ oder „referred journals“ bezeichnet. Manchmal werden nur Veröffentlichungen in solchen Zeitschriften als „echte“ wissenschaftliche

Ergebnisse angesehen – etwas skurril angesichts des Umstands, dass es gar nicht einfach ist festzustellen, ob eine Zeitschrift tatsächlich peer reviewed ist (Pölönen et al. 2019). Irrig ist außerdem die in der Öffentlichkeit häufig geäußerte Annahme, Peer-Review sei eine Garantie für die „Richtigkeit" der Inhalte. Faktisch werden bis auf sehr wenige Ausnahmen (z.B. in der Chemie „Organic Syntheses" (Danheiser 2011)) experimentelle Resultate im Peer-Review nicht repliziert, sondern nur auf Plausibilität überprüft.

Peer-Review wurde vermutlich erstmals vom ersten Sekretär der englischen Royal Society, Henry Oldenburg (1610–1677) praktiziert, ist aber keineswegs seitdem schon Standard bei wissenschaftlichen Zeitschriften. Tatsächlich haben erst im Laufe des neunzehnten und frühen zwanzigsten Jahrhunderts eine Reihe von Fachgesellschaften die Praxis eingeführt, systematisch anonyme Gutachter zu eingereichten Arbeiten zu befragen (Baldwin 2018). Bei der renommierten Zeitschrift *Nature* ist gar erst seit 1973 ein externes Peer-Review obligatorisch (Baldwin 2015).

Der erste Schritt des Verfahrens nach der Einreichung ist die Beurteilung durch den Herausgeber. In diesem Stadium werden einige Beiträge abgelehnt („desk rejection"), weil sie von geringer Qualität sind oder den Rahmen der Zeitschrift sprengen. Sobald sie diese erste Prüfung bestanden haben, werden die Artikel an externe Gutachter weitergeleitet. Das Standardverfahren ist das „blind peer review", bei dem der Autor die Kommentare der Gutachter erhält, die Gutachter selbst aber anonym bleiben. Beim „double-blind peer review" wird der Autor ebenfalls anonymisiert. Dies ist jedoch nicht immer einfach, z.B. wenn Informationen über die Zugehörigkeit des Autors für das Verständnis des Artikels erforderlich sind. Der Grundgedanke dabei ist, den Gutachtern die Möglichkeit zu geben, ihre Meinung offen zu äußern, und im Falle der Doppelblind-Peer-Review jede Voreingenommenheit in Bezug auf Geschlecht, Alter, Ansehen usw. der Autoren zu vermeiden. Die „offene Begutachtung" ist der entgegengesetzte Ansatz, bei dem durch Offenlegung von Gutachtenden und Gutachten versucht wird, einen Bias zu vermeiden (Godlee 2002; Walsh et al. 2000). Beispielsweise wendet der Verlag Copernicus dieses

Verfahren an: Manuskripte werden in der Form der Einreichung veröffentlicht und durchlaufen dann eine öffentliche Begutachtung (jeder kann Kommentare abgeben) sowie eingeladene und eingehendere Begutachtungen. Anschließend wird eine überarbeitete Fassung publiziert. Noch allgemeiner, weil zeitschriftenübergreifend, ist der „PreReview"-Ansatz (Fairhurst 2024). Daneben gibt es noch eine Reihe weiterer Alternativansätze zum klassischen Peer-Review (Tennant et al. 2017).

Der Peer-Review-Prozess für Bücher ist recht heterogen (Goldfinch & Yamamoto 2012). Die externe Begutachtung durch einen (oder sogar mehrere) Gutachter wird häufig durch eine redaktionelle Begutachtung ersetzt, insbesondere bei Sammelbänden, bei denen mehrere Autoren die Kapitel schreiben. Bei Monografien und insbesondere bei Lehrbüchern übernimmt der Verlag das Lektorat durch seine Mitarbeiter.

Anträge auf Forschungsförderung können genauso begutachtet werden wie Zeitschriftenartikel, mit dem Unterschied, dass die Gutachter hier nicht aufgefordert werden, den Text des Antrags zu verbessern, sondern vielmehr ein anderes Versuchsumfeld vorzuschlagen usw. Oft werden Förderanträge von einem Gutachtergremium in einem ganz anderen Rahmen beurteilt: Die Gutachter handeln nicht unabhängig voneinander, sondern diskutieren den Antrag gemeinsam, nachdem sie die Antragsteller befragt haben. Dieses Verfahren umfasst oft eine einzige Sitzung, in der mehrere Vorschläge für dieselbe Ausschreibung geprüft werden. Da ein Finanzhilfeantrag nicht nur den Forschungsvorschlag selbst, sondern auch die Lebensläufe der beteiligten Wissenschaftler mit ihren Veröffentlichungen und Empfehlungsschreiben enthält, wird bei der Begutachtung auch die bisherige Leistung der Antragsteller berücksichtigt. Möglicherweise ist diese anders als der Antragstext sogar der ausschlaggebende Faktor (Simsek et al. 2024).

Eine noch stärkere Verlagerung auf die Bewertung früherer Leistungen findet bei der Bewertung von Forschungsgruppen oder -einrichtungen statt. So unterziehen sich in Deutschland die an der Exzellenzinitiative des Bundes und der Länder beteiligten Universitäten einer strengen Begutachtung durch Wissenschaftler

aus dem Ausland. Das Gleiche gilt für die außeruniversitären Forschungseinrichtungen der Leibniz-Gemeinschaft und der Helmholtz-Gemeinschaft. Sie werden alle sieben Jahre evaluiert, um zu entscheiden, ob sie Mitglied der Leibniz-Gemeinschaft bleiben können oder, im Falle von Helmholtz, in welchem Umfang die Forschungsprogramme finanziert werden.

3. Peer-Review in der Kritik

Peer-Review war lange der Goldstandard in der Wissenschaft und ist es in vielen Zusammenhängen nach wie vor. Es gibt eine Reihe von Argumenten, die für Peer-Review sprechen; dennoch hat die Kritik an Peer-Review in den letzten Jahren zugenommen (Bornmann 2011; Gould 2013; Smith 2006). Einige der kritisierten Aspekte sind (Mittermaier 2023; Ochsner 2022):

- Peer-Reviews sind ineffizient, zeitaufwändig und teuer (Eysenck & Eysenck 1992).
- Veröffentlichungen werden oft verzögert, und innovative und unkonventionelle Ideen können unterdrückt werden (Lee et al. 2013).
- Gutachter:innen könnten die Arbeit absichtlich ablehnen oder zumindest den Prozess verlangsamen und in der Zwischenzeit die Ergebnisse stehlen und auf eigene Faust veröffentlichen (McCook 2016).
- Mehrere Gutachter:innen sind sich selten einig über ihre Empfehlungen, so dass die Reproduzierbarkeit und Zuverlässigkeit schlecht ist (Bornmann et al. 2010).
- Die Empfehlungen der Gutachter:innen sind oft voreingenommen, u.a. bezüglich des Geschlechts der Autor:innen, was bedeutet, dass die Fairness des Prozesses fragwürdig ist (Wennerås & Wold 1997).

4. Bibliometrie

Die Bibliometrie ist ein Instrument zur Ermittlung von objektiven Publikationsdaten, die als Leistungsdaten verwendet werden können. Der Begriff Bibliometrie wurde 1969 geprägt als

eine Bezeichnung für „statistische Bibliographie“ (Pritchard 1969). Eine moderne Definition ist die „Anwendung mathematischer und statistischer Methoden zur Erklärung der Prozesse der schriftlichen Mitteilungen, sowie der Natur und des Entwicklungskurses eines Wissenschaftsgebietes, durch Zählung und Analyse der verschiedenen Aspekte der schriftlichen Kommunikation“ (Gorraiz 2009). Die erste bibliometrische Analyse untersuchte Bücher über die menschliche Anatomie, die zwischen 1550 und 1860 veröffentlicht wurden (Cole & Eales 1917). In der ersten Zitationsanalyse untersuchten die Bibliothekare Gross und Gross die Referenzen in Fachzeitschriften der Chemie (Gross & Gross 1927). Dies ermöglichte es ihnen, eine Rangliste der wichtigsten Fachliteratur der Chemie anhand der Häufigkeit der Zitate zu erstellen, eine frühe Form des Zeitschriftenrankings mit dem Zweck, Bibliotheken bei der Entscheidung über die Beschaffung von Zeitschriften zu helfen. In ihrer Analyse stellten sie eine unregelmäßige Verteilung der Zitate auf die verschiedenen Zeitschriften fest und lieferten damit die Grundlage für das 1934 entwickelte Bradfordsche Gesetz, nach dem sich die wichtigsten wissenschaftlichen Publikationen auf eine Handvoll Kernzeitschriften konzentrieren (Bradford 1934). Brian C. Vickery beschrieb diese „schiefe Verteilung“ mathematisch als *Bradford's Law of Scattering* (Vickery 1948). Bereits 1926 untersuchte der amerikanische Chemiker und Versicherungsmathematiker Alfred James Lotka die Produktivität von Wissenschaftlern und beschrieb die Korrelation zwischen Autoren und Veröffentlichungen, wonach der Publikationsoutput umgekehrt proportional zur Anzahl der Wissenschaftler in einem Fach ist. Mit anderen Worten: nur sehr wenige Autoren haben viele Publikationen, und viele Autoren haben nur wenige (Lotka 1926). Dieser Zusammenhang wird als Lotka'sches Gesetz bezeichnet.

Im Lauf der Zeit wurden neben den Basisgrößen „Zahl der Publikationen“, „Zahl der Zitierungen“ und „Zitationsrate“, als dem Quotienten aus beidem, eine Reihe komplexerer Indikatoren entwickelt. Dazu zählen der Journal Impact Faktor (Tunger 2022), der „Crown Indicator“ (Moed et al. 1995), der eine Feldnormalisierung

enthält und somit verschiedene Wissenschaftsgebiete vergleichbar macht, sowie der h-Index (Hirsch 2005), der quantitative und qualitative Aspekte in einer einzigen Zahl vereinigen möchte, tatsächlich aber proportional zur Wurzel aus der Gesamtzahl der Zitierungen ist (Krattenthaler 2021).

5. Bibliometrie in der Kritik

Folgende Kritikpunkte werden gegenüber der Bibliometrie vorgebracht; die Auflistung ist nicht abschließend:

- Bibliometrie bewertet den Impact einer Publikation, indem sie die Zahl der Zitierungen misst. Dies ignoriert den Umstand, dass es ganz unterschiedliche Gründe für eine Zitierung geben kann, bis hin zur Ablehnung der zitierten Aussage (Cronin 1981).
- Da umgekehrt wissenschaftliche Publikationen andere Wissenschaftler:innen und deren Publikationen auch jenseits von Zitierungen stark beeinflussen können, wurde schon die Grundlage bibliometrischer Analysen an sich in Frage gestellt (MacRoberts & MacRoberts 1986).
- Zitationen haben naturgemäß einen ziemlichen zeitlichen Versatz: Meist wird bei bibliometrischen Analysen ein dreijähriges Zitationsfenster angelegt, was bedeutet, dass Publikationen des Jahres 2024 erst ab dem Jahr 2027 bewertet werden.
- Bibliometrie ist Grundlage vieler zweifelhafter Rankings (van Raan 2005).
- Bibliometrische Indikatoren sind manipulierbar, beispielsweise durch ein Übermaß an Selbstzitierungen, durch Zitationskartelle oder durch erzwungene Zitierungen auf die gleiche Zeitschrift (Böhme & Tesch 2017).
- Es wird immer wieder der Fehler gemacht, bibliometrische Indikatoren ohne Normierung zu verwenden (Tunger, Ahn, Clermont, Krolak & Meier 2022). Beispielsweise kann der h-Index zwar als Maß für die Exzellenz von Wissenschaftler:innen verwendet werden. Er hängt aber nicht nur von der Qualität der Arbeit, sondern auch vom Lebensalter und vom Fach ab (Fraumann & Mutz 2022).
- In manchen Fällen werden anhand von bibliometrischen Indikatoren Aussagen getroffen, die keinerlei Grundlage im Indikator selbst

haben. Berühmtestes Beispiel ist der Journal Impact Faktor (IF), der ein Maß für die Zitationshäufigkeit einer Zeitschrift ist, der aber keine Aussage über einzelne Artikel oder über Personen erlaubt (Tunger 2022).
- Wissenschaftler:innen tendieren dazu, einer Bewertung durch ihresgleichen, wie sie im Peer-Review stattfindet, größeres Vertrauen entgegenzubringen als einer bibliometrischen Analyse, die oftmals durch die Bibliothek oder das Forschungsdezernat durchgeführt wird (Rühle 2009).

6. Altmetrics[3]

Parallel zur Entstehung und Entwicklung des Internets hat sich auch die Wissenschaftskommunikation weiterentwickelt. Sie findet nun auch z.B. in Form von Blogs und sozialen Netzwerken statt, was neue Möglichkeiten zur Verbreitung und zum Austausch wissenschaftlicher Erkenntnisse bietet (Moed 2016; Sugimoto et al. 2017). In der Konsequenz scheint die „klassische" Impact-Messung über Zitationen v.a. in wissenschaftlichen Zeitschriften als alleinige Methode nicht mehr ausreichend. Neue Wege des wissenschaftlichen Austauschs und der Publikation erhöhen die Datenbasis um ein Vielfaches, traditionelle bibliometrische Methoden zur Messung geraten dort an ihre Grenzen (Cronin & Sugimoto 2014).

Der Begriff „Altmetrics", ein Kofferwort aus „alternative" und „metrics", wurde 2010 von Jason Priem in einem Tweet geprägt: „I like the term #articlelevelmetrics, but it fails to imply *diversity* of measures. Lately, I'm liking #altmetrics"[4]. Im gleichen Jahr war er einer der Autoren des „Altmetrics manifesto" (Priem et al. 2010). Definiert wird Altmetrics als „the study and use of scholarly impact measures based on activity in online tools and environments." (Priem 2014) Altmetrics entstehen durch Aktivitäten in sozialen Medien, beispielsweise durch Erwähnungen in Blogs, Tweets auf Twitter, Likes auf Facebook oder das Abspeichern in Literaturverwaltungsprogrammen. Allerdings gibt es keinen „kanonischen" Satz von Plattformen/Diensten, die „Altmetrics" bilden. Beispielsweise wurde vorgeschlagen (Priem et al. 2012):

- „social media like Twitter and Facebook
- online reference managers like CiteULike, Zotero, and Mendeley
- collaborative encyclopedias like Wikipedia
- blogs, both scholarly and general-audience
- scholarly social networks, like ResearchGate or Academia.edu
- conference organization sites like Lanyrd.com“

Aus heutiger Perspektive betrachtet: Viele Personen und Organisationen haben Twitter (jetzt: X) nach der Übernahme durch Elon Musk verlassen und die Bedeutung von Facebook für Wissenschaftskommunikation kann stark hinterfragt werden. Die Konferenzplattform Lanyrd existiert seit 2018 nicht mehr.

Das Unternehmen Altmetric Ltd. (Digital Science), das auf Personen- und Artikelebene den „Altmetric Attention Score“ (AAS) und daraus den „Altmetric Donut“ erstellt, erhebt derzeit ein Set von 15 Indikatoren[5] (Abb. 6.1). Von den genannten Quellen liefern Sina Weibo, Syllabi, Google+ und Pinterest aktuell aber keine Daten mehr oder existieren nicht mehr (Altmetric Ltd. 2024a) – ein weiterer Hinweis auf die Instabilität von „Altmetrics“.[6] Ein konkurrierender Indikator sind die PlumX Metrics von Plum Analytics. Sie enthalten sehr viele Metriken aus den Bereichen Zitierung, Nutzung, Markierung, Erwähnung und soziale Medien (Elsevier 2024). Wie auch Plum Analytics selbst gehören viele von ihnen zum Elsevier-Konzern.

Die Annahme liegt nahe, dass Altmetrics (auch) etwas anderes messen als bibliometrische Indikatoren, nämlich den gesellschaftlichen Einfluss, den wissenschaftliche Arbeit hat

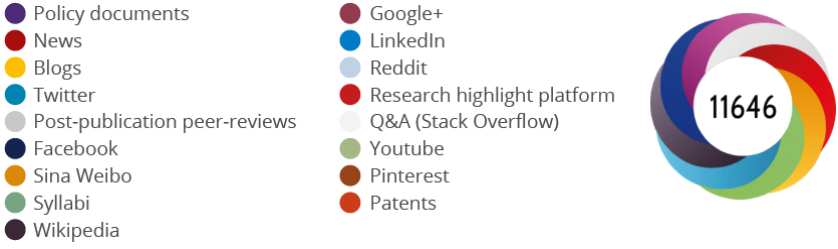


Abbildung 6.1: Altmetric-Donut (Altmetric Ltd. 2024b).

(Bornmann 2014; Haustein 2016). Es gibt einige derartige Untersuchungen, zum Teil mit dem Ergebnis, dass eine Disziplin gesellschaftlich isoliert ist (Conix et al. 2022). Ein Gegenstand noch intensiverer Forschung war die Korrelation zwischen Bibliometrie und Altmetrics (Costas et al. 2015). Eine Metastudie (Stephen & Stahlschmidt 2024) wertete die in 111 Studien zu 13 verschiedenen altmetrischen Indikatoren gefundenen Korrelationen mit Zitationen aus. Deutlich positive Korrelationen gab es mit der Nennung in Mendeley (ein Literaturverwaltungsprogramm), Nutzungsstatistiken, der Präsenz auf ResearchGate und der Twitter-Aktivität.

7. Altmetrics in der Kritik

Die nachfolgend genannten Kritikpunkte an den Altmetrics sind zahlenmäßig geringer als die entsprechende Auflistung bei Bibliometrie, ihr Gewicht ist aber umso größer:

- Zusammensetzung und Gewichtung der altmetrischen Faktoren ist willkürlich und ohne valide Begründung.
- Anders als z.B. bei einem Börsenindex kann man bei Wegfall eines altmetrischen Indikators und Ersatz durch einen anderen nicht sicherstellen, dass Werte vor und nach der Änderung miteinander vergleichbar sind. Altmetrische Zeitreihen sind somit weitgehend sinnlos.
- Altmetrics sind in noch viel größerem Maß manipulationsanfällig als klassische bibliometrische Indikatoren, z.B. durch Bots und „Gaming“ (Aide 2013; Biagioli & Lippman 2020; Kulczycki 2023).
- Die Aussagekraft vieler altmetrischer Indikatoren ist gering. Oft liefern sie kaum zusätzliche Information zu klassischen bibliometrischen Indikatoren. (Bornmann & Haunschild 2018; Tunger, Ahn, Clermont, Krolak, Langner & Mittermaier 2022).
- Insgesamt wird von der Verwendung von Altmetrics in formellen Evaluierungen klar abgeraten: „This is a potentially disastrous error relating to practical applications of altmetrics. Nearly all alternative indicators lack quality control and are relatively easy to manipulate accidentally or deliberately.“ (Thelwall 2021)

8. Vergleich von Peer-Review mit bibliometrischen und altmetrischen Ansätzen

In Anbetracht der Kritik an Peer-Review (siehe oben) sind Alternativen prüfenswert. Im Falle von Förderanträgen wurde das „Peer-Reviewed-Productivity-Formel-System" (Roy 1985) vorgeschlagen, es kam aber offenbar nicht produktiv in den Einsatz. Eine andere Möglichkeit ist die Verwendung bibliometrischer oder altmetrischer Indikatoren anstelle von Peer-Reviews. Dieser Ansatz war bereits Gegenstand zahlreicher Untersuchungen, für Übersichten vgl. z.B. (Aksnes & Taxt 2004; Bornmann 2011; Gallo & Glisson 2018; Vieira & Gomes 2018). Nachstehend werden verschiedene Begutachtungsszenarien bzw. Anwendungsfälle dargestellt.

8.1 Peer-Review vs. Journal Impact Factor

Es ist allgemein anerkannt, dass es eine Hierarchie der Zeitschriften gibt, manchmal veranschaulicht mit dem Bild einer Pyramide: Es gibt einige wenige hochrangige Zeitschriften (definiert als solche durch ihren Journal Impact Factor (JIF) gemäß den von Clarivate Analytics veröffentlichten Journal Citation Reports), gefolgt von einer größeren Anzahl von Zeitschriften mit mittlerem Impact und einer noch größeren Anzahl von Zeitschriften mit geringem Impact (McKiernan et al. 2019). „Um in der Wissenschaft erfolgreich zu sein, muss man diese Pyramide erklimmen: zumindest im akademischen Bereich ist die Veröffentlichung in den prestigeträchtigeren Zeitschriften der Schlüssel zum beruflichen Aufstieg." (Jennings 2006) Dieser Annahme zufolge reichen Autoren ihr Manuskript bei einer der am höchsten eingestuften Zeitschriften ein und steigen im Falle einer Ablehnung in der Pyramide nach unten und reichen bei Zeitschriften mit einem niedrigeren JIF ein. Bornmann (2011) zitiert ein Dutzend Studien, die die Qualität der ablehnenden und der anschließend akzeptierenden Zeitschriften anhand des JIF verglichen haben. Bis zu 70 Prozent der abgelehnten Manuskripte konnten als später in einer Zeitschrift mit einem höheren JIF veröffentlicht nachverfolgt werden.

Daher „wechseln Autoren nach einer Ablehnung nicht notwendigerweise von ‚führenden' Zeitschriften zu weniger angesehenen Zeitschriften" (Weller 2001).

8.2 Peer-Review vs. Zitationsanalyse

Sir Theodore F. Fox, ehemaliger Herausgeber von *The Lancet*, war eher skeptisch, was den Vorhersagewert redaktioneller Entscheidungen angeht: „Wenn ich die Beiträge der Woche in zwei Stapel aufteile – einen, den wir veröffentlichen werden, und den anderen, den wir zurückschicken werden –, dann frage ich mich, ob es für die Zeitschrift oder ihre Leser einen wirklichen Unterschied machen würde, wenn ich den einen Stapel gegen den anderen austausche." (Fox 1965) Evidenzbasierte Untersuchungen zeigen jedoch, dass diese Entscheidungen bei einer Reihe von Zeitschriften erhebliche Auswirkungen haben: (Bornmann & Daniel 2008) untersuchten 878 von der *Angewandten Chemie – International Edition* angenommene und 959 abgelehnte Artikel, die dann anderswo veröffentlicht wurden. Ihre Ergebnisse zeigen, dass die Annahme durch die *Angewandte Chemie* die erwartete Anzahl von Zitationen um bis zu 50 % erhöht. Ähnliche Ergebnisse wurden in Untersuchungen zum *American Journal of Neuroradiology* (McDonald et al. 2007; 2009), *Cardiovascular Research* (Opthof et al. 2000) und zur Publikationsplattform *F1000* (Bornmann & Leydesdorff 2013) erzielt. Breiter angelegte Untersuchungen von (Cicchetti 1991) und (Benda & Engels 2011) haben diese Ergebnisse bestätigt. Die Zitationsanalyse deutet also darauf hin, dass bei dem Peer-Review tendenziell der „bessere Haufen" von Manuskripten ausgewählt wird, der dann auch häufiger gelesen und zitiert wird.

8.3 Förderentscheidungen

In einigen Studien wurden Peer-Reviews zu Fördermitteln mit dem Ziel analysiert, festzustellen, ob Antragsteller, die eine Förderung erhalten haben, häufiger zitiert wurden als Antragsteller

ohne Förderung: (van Leeuwen & Moed 2012) analysierten die Korrelation zwischen der Finanzierung durch drei Fördergremien in den Niederlanden. Erfolgreiche Antragsteller erzielten tendenziell einen höheren Zitationsimpakt auf internationaler Ebene als diejenigen, deren Anträge abgelehnt wurden. Auch einige andere Studien (Bornmann & Daniel 2006; Bornmann & Daniel 2008; Cabezas-Clavijo et al. 2013) ergaben eine positive Korrelation zwischen der Begutachtung von Stipendien und der Zitationswirkung. Andere Untersuchungen zeigten jedoch keine oder nur eine geringe Korrelation zwischen erfolgreichen Förderanträgen (Emmy Noether-Programm der DFG, Schwedische Stiftung für Strategische Forschung und Rat für Sozialwissenschaftliche Forschung der Niederlande) und dem anschließenden Zitationserfolg (Hornbostel et al. 2009; Melin & Danell 2006; van Besselaar & Leydesdorff 2009). Die Ergebnisse des Vergleichs von Peer-Review und bibliometrischen Indikatoren sind also gemischt. Darüber hinaus wurde auch die Frage gestellt, ob die Auswahl der Antragsteller eine sich selbst erfüllende Prophezeiung sein könnte: „Sie belohnen frühere Exzellenz [...], aber sie bieten den erfolgreichen Antragstellern auch Ressourcen, die sie in die Lage versetzen könnten, in ihrer zukünftigen Karriere exzellente wissenschaftliche Arbeit zu leisten. Wenn eine Studie signifikante Leistungsvorteile für angenommene, aber nicht für abgelehnte Bewerber feststellt, kann man argumentieren [...], dass die Förderorganisation den Stipendiaten einen solchen Vorteil in Bezug auf Ausbildung, Prestige, Selbstvertrauen usw. verschafft, dass sie später aufgrund des Stipendiums zu überlegenen Wissenschaftlern werden, nicht weil sie zum Zeitpunkt der Bewerbung besonders vielversprechend waren. Anstatt die ‚besten' Wissenschaftler auszuwählen, könnte der Auswahlausschuss sie nach dieser Auffassung schaffen." (Bornmann 2011)

8.4 Berufungsverfahren

Vieira et al. (2014) analysierten 27 Einstellungsverfahren mit 174 Kandidat:innen für Professuren im STM-Bereich, die zwischen

2007 und 2011 an sechs portugiesischen Universitäten stattfanden. Wenn zwei beliebige Bewerber:innen in einem bestimmten Einstellungsverfahren anhand einer Kombination von zwei bibliometrischen Indikatoren (einer h-Index-Variante und dem Prozentsatz hoch zitierter Dokumente) verglichen wurden, konnte das Ergebnis der Einstufung dieser beiden Bewerber durch die Fachkollegen in 75 % der Fälle vorhergesagt werden. Jensen et al. (2009) untersuchten die Korrelation zwischen fünf verschiedenen bibliometrischen Indikatoren und den Ergebnissen eines Peer-Review-Verfahrens für die Beförderung von etwa 600 Forschenden am französischen Centre Nationale de la Recherche Scientifique (CNRS). Die Autoren stellten fest, dass „kein einzelner Indikator der beste Prädiktor für alle Disziplinen ist. Insgesamt liefert der Hirsch-Index h die am wenigsten schlechten Korrelationen, gefolgt von der Anzahl der veröffentlichten Arbeiten. Es ist jedoch wichtig zu erkennen, dass selbst h nur die Hälfte der tatsächlichen Beförderungen wiedergeben kann. Die Anzahl der Zitierungen oder die durchschnittliche Anzahl der Zitierungen pro Arbeit sind definitiv keine guten Prädiktoren für eine Beförderung." (ebd.)

8.5 Bewertung von Forschungsgruppen

Es gibt mehrere Studien, die das Peer-Review von Forschungsgruppen mit bibliometrischen Indikatoren vergleichen. Im Allgemeinen fanden diese Studien eine (manchmal schwache) positive Korrelation. Mike Thelwall erklärt die unvollkommenen Korrelationen zwischen bibliometrischen Indikatoren und Peer-Review (teilweise) durch die Variation in den qualitativen Peer-Bewertungen (Thelwall et al. 2015). Anthony van Raan et al. führten eine Reihe von Untersuchungen in diesem Bereich durch: Nederhof & van Raan (1993) analysierten die Beziehung zwischen bibliometrischen Indikatoren und Peer-Review für sechs Forschungsgruppen in den Wirtschaftswissenschaften. Peer-Review und bibliometrische Ergebnisse stimmten im Allgemeinen überein. Für die Niederlande konnte die Korrelation zwischen

verschiedenen bibliometrischen Indikatoren und den Ergebnissen des Peer-Review durch Expertengremien von Physikern gezeigt werden (Rinia et al. 1998). Hierbei wurden 56 Forschungsprogramme mit ca. 5.000 Veröffentlichungen und 50.000 Zitierungen untersucht. Die stärkste Korrelation bestand zwischen bibliometrischen Indikatoren und der Beurteilung der Forscher und des Forschungsteams. Später untersuchte van Raan die Korrelation zwischen bibliometrischen Standardindikatoren und der Beurteilung durch Fachkollegen für 147 chemische Forschungsgruppen in den Niederlanden (van Raan 2006). Er fand heraus, dass sowohl der h-Index als auch der „Crown Indicator" sehr gut zwischen den Gruppen, die eine Bewertung von 3 („zufriedenstellend") erhielten, und den Gruppen, die mit 4 („gut") und 5 („ausgezeichnet") bewertet wurden, unterscheiden.

8.6 Nationale Forschungsbewertungen

Eine vergleichende Analyse von Zitationsindikatoren und Peer-Review im Rahmen der italienischen Forschungsbewertung Valutazione Triennale della Ricerca (VTR) ergab eine signifikante Korrelation zwischen den beiden Ansätzen (Ancaiani et al. 2015), wobei diese Schlussfolgerung später aus methodischen Gründen in Frage gestellt wurde (Baccini & Nicolao 2016). Die letztgenannte Studie kam zu dem Schluss, dass Bibliometrie und Peer-Review nicht zu ähnlichen Ergebnissen führen. Zwei weitere Studien identifizieren eine positive Korrelation zwischen Peer-Entscheidungen und dem Impact-Faktor der Zeitschriften, in denen die Dokumente im italienischen VTR veröffentlicht wurden, wobei sie deutliche Unterschiede zwischen den wissenschaftlichen Bereichen feststellten (Abramo et al. 2009; Franceschet & Costantini 2011).

Eine ganze Reihe von Untersuchungen wurde mit Daten aus dem britischen Research Assessment Exercise (RAE) und seinem Nachfolger, dem Research Excellence Framework (REF), durchgeführt. Beim RAE 1992 gab es für drei Fachbereiche (Anatomie, Archäologie und Genetik) statistisch signifikante Korrelationen zwischen der Gesamtzahl der erhaltenen Zitate

bzw. der durchschnittlichen Anzahl der Zitate pro Mitarbeiter und der RAE-Punktzahl (Oppenheim 1997). Ähnliche Ergebnisse wurden anhand der RAE-Daten von 2001 für den Bereich der Archäologie ermittelt (Norris & Oppenheim 2003). Während einige Experten die Schlussfolgerungen kritisierten (Warner 2000), replizierten andere die Ergebnisse für die Bereiche Politikwissenschaft und Chemie. Es gebe jedoch „kein einziges Modell, das für alle wissenschaftlichen und nicht-wissenschaftlichen Disziplinen gilt. Jeder metrische Ansatz zur Leistungsbewertung muss eine disziplinspezifische Reihe von Indikatoren verwenden." (Butler & Mcallister 2011) Eine Überprüfung der bibliometrischen Daten des RAE 2001 kam zu demselben Ergebnis. Es gab enorme Unterschiede zwischen den Wissenschaftsbereichen, und selbst innerhalb eines Bereichs mit einer statistisch signifikanten Korrelation zwischen Zitation und RAE-Punktzahl, wie z.B. der Chemie, war die Rangfolge der Institutionen, die sich aus den beiden Messungen ergab, bei weitem nicht identisch (Mahdi, D'Este & Neely 2008). Die nachfolgenden RAE- und REF-Bewertungen waren Gegenstand weiterer Untersuchungen; für einen Überblick siehe (Traag & Waltman 2019). Bei der REF 2021 wurden auch altmetrische Indikatoren mit Peer-Review verglichen. Die Korrelation mit Peer-Review-Ergebnissen ist vorhanden, aber weniger stark als bei Zitationsindikatoren. Im Allgemeinen ist die Korrelation im STM-Bereich größer als in den Geisteswissenschaften (Thelwall et al. 2023).

9. Schlussfolgerungen

Insgesamt wurde in den meisten Vergleichsstudien eine mäßig positive Korrelation zwischen Peer-Review und bibliometrischen Indikatoren festgestellt, aber die ermittelten Korrelationen waren nicht stark und variierten zwischen den Studien. Unter anderem hängt die Stärke der Korrelationen vom wissenschaftlichen Bereich, den bibliometrischen Indikatoren und dem Gegenstand der Begutachtung ab.

- Die Ergebnisse von Studien, die sich auf Förderentscheidungen konzentrieren, sind uneinheitlich. Während eine Reihe von Studien eine positive Korrelation zwischen der Begutachtung von Finanzhilfen und der Zitierwirkung ergab, zeigten andere Studien keine oder nur eine geringe Korrelation zwischen dem Erfolg bei Finanzhilfeanträgen und der späteren Zitierwirkung.
- Auch die Ergebnisse von Untersuchungen zu Tenure-Entscheidungen sind nicht überzeugend. Bestenfalls können bibliometrische Indikatoren in 75 % der Fälle die richtige Einstufung von zwei Bewerbern vorhersagen, was nur halbwegs zwischen dem tatsächlichen Ergebnis (100 %) und einer Zufallsentscheidung (50 %) liegt. Überprüfungen der Bewertungen von Forschungsgruppen ergaben im Allgemeinen eine bessere Übereinstimmung zwischen Peer-Review und Bibliometrie, oft jedoch abhängig vom wissenschaftlichen Bereich und dem betreffenden Indikator.
- Die Untersuchungen zu den nationalen Forschungsbewertungen ergaben ähnliche Ergebnisse wie die Studien zu den Bewertungen der Forschungsgruppen. Als Faustregel gilt, dass die Korrelationen im Bereich Wissenschaft, Technologie und Medizin (STM) besser sind als in den Sozial- und Geisteswissenschaften, und dass die Korrelationen für feldnormierte Indikatoren besser sind als für Basisindikatoren wie die Anzahl der Zitate (Marx & Bornmann 2015).

Daher gibt es im Allgemeinen wenig empirische Unterstützung für die Hypothese, dass die Bibliometrie dieselben Aspekte der Wirkung oder Forschungsqualität widerspiegelt wie Peer-Review. Das Ausmaß, in dem die Korrelation zwischen den beiden Ansätzen als ausreichend angesehen wird, hängt jedoch von der Art und den Zielen der Bewertung ab. Die Aussage von (Abramo & D'Angelo 2011) zu nationalen Forschungsbewertungen könnte auch für die Bewertung von Forschungsgruppen gelten: „Unter der Annahme, dass es nicht die eine unfehlbare Evaluierungsmethode gibt, vertreten die Autoren die Position, dass für die Natur- und Formalwissenschaften die bibliometrische Methodik bei weitem dem informierten Peer-Review vorzuziehen ist." Diese These mag auf den ersten Blick gewagt erscheinen, ist aber aus folgenden Gründen je nach Anwendungsfall gerechtfertigt:

- Das Peer-Review-Verfahren ist bei weitem nicht perfekt reproduzierbar (Cole et al. 1981). Daher bedeutet eine Abweichung der Ergebnisse eines bibliometrischen Ansatzes von einer Peer-Entscheidung nicht unbedingt, dass der bibliometrische Ansatz zu einem „falschen" Ergebnis geführt hat.
- Peer-Reviews sind sehr arbeitsintensiv und daher teuer. Dies ist z.B. bei Berufungsentscheidungen zu rechtfertigen, muss aber bei der Erstellung von Rankings in Frage gestellt werden.
- Während es gute Gründe gibt, bei der Beurteilung von Einzelpersonen (z.B. Tenure, Grant) mit bibliometrischen Ansätzen als einziger Methode zurückhaltend zu sein, sieht es bei der Beurteilung von Gruppen anders aus, und hier könnten bibliometrische Analysen das Peer-Review ersetzen

Hinsichtlich der Verwendung von Altmetrics ist der Forderung von Butler et al. beizupflichten: „Wir schlagen vor, dass Altmetrics, die bei der Forschungsevaluierung eingesetzt werden, Teil eines fundierten Peer-Review-Prozesses sein sollten, so wie die traditionellen Metriken. Darüber hinaus dürfen die auf Altmetrics basierenden Ergebnisse nicht zu einer direkten Entscheidungsfindung über die Forschung führen, sondern sollten stattdessen dazu dienen, Experten bei der Entscheidungsfindung zu unterstützen. Schließlich sollten sich traditionelle und alternative Metriken im Peer-Review-Verfahren gegenseitig ergänzen und nicht ersetzen." (Butler et al. 2017)

Im Ergebnis ist also Peer-Review das Mittel der Wahl bei kleinskaligen Anwendungen (Manuskripte, einzelne Personen). Bei Forschergruppen ist die Unterstützung durch Bibliometrie sinnvoll, bei nationalen Auswertungen ist dies unabdingbar. Altmetrics eignen sich nicht zur Forschungsevaluierung. Ihr Vorteil liegt aber darin, dass eine gewisse Vorhersagekraft für den späteren Zitationserfolg von Publikationen besteht.

Anmerkungen

[1] Dieser Beitrag beruht auf dem Buchbeitrag (Mittermaier 2022) und dem Vortrag (Mittermaier 2023).

[2] In einem weiteren Sinne wird der Begriff „Peer-Review" auch verwendet, um z.B. die Lehre (Chism 2007) oder medizinisches Personal (Chop & Eberlein-Gonska 2012) zu bewerten. Diese Aspekte werden hier nicht diskutiert.

[3] Die deutschen Begriffe „Altmetrik" und „Altmetrie" existieren, sind aber völlig ungebräuchlich. Praktisch verwendet wird jedoch das Adjektiv „altmetrisch".

[4] https://x.com/jasonpriem/status/25844968813 (Zugriff am 15.08.2024).

[5] Die Quellen werden ohne nachvollziehbare Begründung nicht gleich gewichtet, https://help.altmetric.com/support/solutions/articles/6000233311-how-is-the-altmetric-attention-score-calculated, was Anlass zur Kritik gibt.

[6] Diese Abbildung stammt ausweislich ihrer URL, https://wordpress-uploads-production.s3.amazonaws.com/uploads/2022/10/Screenshot-2018-10-30-at-16.30.13.png, wohl aus dem Jahr 2018, wird aber so auf der aktuellen Webseite verwendet.

Literatur

Abramo, G.; D'Angelo, C. A. (2011) „Evaluating research: from informed peer review to bibliometrics", *Scientometrics*, Vol. 87, No. 3, S. 499–514 [Online]. DOI: https://doi.org/10.1007/s11192-011-0352-7.

Abramo, G.; D'Angelo, C. A.; Caprasecca, A. (2009) „Allocative efficiency in public research funding: Can bibliometrics help?", *Research Policy*, Vol. 38, No. 1, S. 206–215 [Online]. DOI: https://doi.org/10.1016/j.respol.2008.11.001.

Aide, E. (2013) *Gaming altmetrics* [Online]. Verfügbar unter https://www.altmetric.com/blog/gaming-altmetrics/ (Abgerufen am 20 August 2024).

Aksnes, D. W.; Taxt, R. E. (2004) „Peer reviews and bibliometric indicators: a comparative study at a Norwegian university", *Research Evaluation*, Vol. 13, No. 1, S. 33–41 [Online]. DOI: https://doi.org/10.3152/147154404781776563.

Altmetric Ltd. (2024a) *Our Sources* [Online]. Verfügbar unter https://www.altmetric.com/about-us/our-data/our-sources/ (Abgerufen am 20 August 2024).

Altmetric Ltd. (2024b) *The donut and Altmetric Attention Score* [Online]. Verfügbar unter https://www.altmetric.com/about-us/our-data/donut-and-altmetric-attention-score/ (Abgerufen am 20 August 2024).

Ancaiani, A.; Anfossi, A. F.; Barbara, A.; Benedetto, S.; Blasi, B.; Carletti, V.; Cicero, T.; Ciolfi, A.; Costa, F.; Colizza, G.; Costantini, M.; Di Cristina, F.; Ferrara, A.; Lacatena, R. M.; Malgarini, M.; Mazzotta, I.; Nappi, C. A.; Romagnosi, S.; Sileoni, S. (2015) „Evaluating scientific research in Italy: The 2004–10 research evaluation exercise", *Research Evaluation*, Vol. 24, No. 3, S. 242–255 [Online]. DOI: https://doi.org/10.1093/reseval/rvv008.

Baccini, A.; Nicolao, G. de (2016) „Do they agree? Bibliometric evaluation versus informed peer review in the Italian research assessment exercise", *Scientometrics*, Vol. 108, No. 3, S. 1651–1671 [Online]. DOI: https://doi.org/10.1007/s11192-016-1929-y.

Baldwin, M. (2015) „Credibility, Peer Review, and Nature, 1945–1990", *Notes and records of the Royal Society of London*, Vol. 69, No. 3, S. 337–352 [Online]. DOI: https://doi.org/10.1098/rsnr.2015.0029.

Baldwin, M. (2018) „Scientific Autonomy, Public Accountability, and the Rise of ‚Peer Review' in the Cold War United States", *Isis*, Vol. 109, No. 3, S. 538–558 [Online]. DOI: https://doi.org/10.1086/700070.

Ball, R (Hg.) (2022) *Handbook bibliometrics* [Online], Berlin, Boston, De Gruyter Saur. Verfügbar unter https://doi.org/10.1515/9783110646610.

Benda, W. G.; Engels, T. C. (2011) „The predictive validity of peer review: A selective review of the judgmental forecasting qualities of peers, and implications for innovation in science", *International Journal of Forecasting*, Vol. 27, No. 1, S. 166–182 [Online]. DOI: https://doi.org/10.1016/j.ijforecast.2010.03.003.

Biagioli, M; Lippman, A (Hg.) (2020) *Gaming the metrics: Misconduct and manipulation in academic research*, Cambridge, Mass., MIT Press.

Böhme, U.; Tesch, S. (2017) „Die dunkle Seite der Bibliometrie", *Nachrichten aus der Chemie*, Vol. 65, No. 10, S. 1024–1027 [Online]. DOI: https://doi.org/10.1002/nadc.20174065326.

Bornmann, L. (2011) „Scientific peer review", *Annual Review of Information Science and Technology*, Vol. 45, No. 1, S. 197–245 [Online]. DOI: https://doi.org/10.1002/aris.2011.1440450112.

Bornmann, L. (2014) „Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics",

Journal of Informetrics, Vol. 8, No. 4, S. 895–903 [Online]. DOI: https://doi.org/10.1016/j.joi.2014.09.005.

Bornmann, L.; Daniel, H.-D. (2006) „Selecting scientific excellence through committee peer review – A citation analysis of publications previously published to approval or rejection of post-doctoral research fellowship applicants", *Scientometrics*, Vol. 68, No. 3, S. 427–440 [Online]. DOI: https://doi.org/10.1007/s11192-006-0121-1.

Bornmann, L.; Daniel, H.-D. (2008) „Selecting manuscripts for a high-impact journal through peer review: A citation analysis of communications that were accepted by Angewandte Chemie International Edition, or rejected but published elsewhere", *Journal of the American Society for Information Science and Technology*, Vol. 59, No. 11, S. 1841–1852 [Online]. DOI: https://doi.org/10.1002/asi.20901.

Bornmann, L.; Haunschild, R. (2018) „Do altmetrics correlate with the quality of papers? A large-scale empirical study based on F1000Prime data", *PloS one*, Vol. 13, No. 5, e0197133 [Online]. DOI: https://doi.org/10.1371/journal.pone.0197133.

Bornmann, L.; Leydesdorff, L. (2013) „The validation of (advanced) bibliometric indicators through peer assessments: A comparative study using data from InCites and F1000", *Journal of Informetrics*, Vol. 7, No. 2, S. 286–291 [Online]. DOI: https://doi.org/10.1016/j.joi.2012.12.003.

Bornmann, L.; Mutz, R.; Daniel, H.-D. (2010) „A reliability-generalization study of journal peer reviews: a multilevel meta-analysis of inter-rater reliability and its determinants", *PloS one*, Vol. 5, No. 12, e14331 [Online]. DOI: https://doi.org/10.1371/journal.pone.0014331.

Bradford, S. C. (1934) „Sources of Information on Scientific Subjects", *Engineering: An Illustrated Weekly Journal*, Vol. 137, S. 85–86.

Butler, J. S.; Kaye, I. D.; Sebastian, A. S., Wagner, S. C., Morrissey, P. B., Schroeder, G. D., Kepler, C. K.; Vaccaro, A. R. (2017) „The Evolution of Current Research Impact Metrics: From Bibliometrics to Altmetrics?", *Clinical spine surgery*, Vol. 30, No. 5, S. 226–228 [Online]. DOI: https://doi.org/10.1097/BSD.0000000000000531.

Butler, L.; Mcallister, I. (2011) „Evaluating University Research Performance Using Metrics", *European Political Science*, Vol. 10, No. 1, S. 44–58 [Online]. DOI: https://doi.org/10.1057/eps.2010.13.

Cabezas-Clavijo, A.; Robinson-García, N.; Escabias, M.; Jiménez-Contreras, E. (2013) „Reviewers' ratings and bibliometric indicators: hand in hand when assessing over research proposals?", *PloS one*, Vol. 8, No. 6, e68258 [Online]. DOI: https://doi.org/10.1371/journal.pone.0068258.

Chism, N. (2007) *Peer review of teaching: A sourcebook*, 2. Aufl., Bolton, MA, Anker.

Chop, I.; Eberlein-Gonska, M. (2012) „Übersichtsartikel zum Peer Review Verfahren und seine Einordnung in der Medizin", *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, Vol. 106, No. 8, S. 547–552 [Online]. DOI: https://doi.org/10.1016/j.zefq.2012.08.017.

Cicchetti, D. V. (1991) „The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation", *Behavioral and Brain Sciences*, Vol. 14, No. 1, S. 119–135 [Online]. DOI: https://doi.org/10.1017/S0140525X00065675.

Cole, F. J.; Eales, N. B. (1917) „The history of comparative anatomy: Part I.-A statistical analysis of the literature", *Science Progress*, Vol. 11, No. 44, S. 578–596.

Cole, S.; Cole, J. R.; Simon, G. A. (1981) „Chance and consensus in peer review", *Science*, Vol. 214, No. 4523, S. 881–886 [Online]. DOI: https://doi.org/10.1126/science.7302566.

Conix, S.; Lemeire, O.; Chi, P.-S. (2022) „The public relevance of philosophy", *Synthese*, Vol. 200, No. 1 [Online]. DOI: https://doi.org/10.1007/s11229-022-03546-9.

Costas, R.; Zahedi, Z.; Wouters, P. (2015) „Do ‚altmetrics' correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective", *Journal of the Association for Information Science and Technology*, Vol. 66, No. 10, S. 2003–2019 [Online]. DOI: https://doi.org/10.1002/asi.23309.

Cronin, B; Sugimoto, CR (Hg.) (2014) *Beyond bibliometrics: Harnessing multidimensional indicators of scholarly impact*, Cambridge, Massachusetts, MIT Press.

Cronin, B. (1981) „The need for a theory of citing", *Journal of Documentation*, Vol. 37, No. 1, S. 16–24 [Online]. DOI: https://doi.org/10.1108/eb026703.

Danheiser, R. L. (2011) „Organic Syntheses: The ‚Gold Standard' in Experimental Synthetic Organic Chemistry", *Organic Syntheses*, Vol. 88, S. 1–3 [Online]. DOI: https://doi.org/10.15227/orgsyn.088.0001.

Derrick, G. (2019) *The evaluators' eye: Impact assessment and academic peer review*, Cham, Switzerland, Palgrave Macmillan.

Elsevier (2024) *PlumX Metrics* [Online]. Verfügbar unter https://www.elsevier.com/insights/metrics/plumx (Abgerufen am 20 August 2024).

Eysenck, H. J.; Eysenck, S. (1992) „Peer review: Advice to referees and contributors", *Personality and Individual Differences*, Vol. 13, No. 4, S. 393–399 [Online]. DOI: https://doi.org/10.1016/0191-8869(92)90066-X.

Fairhurst, V. (2024) *Transforming Scholarly Publishing: A Community-Led Approach to Peer Review with PREreview* [Online]. Verfügbar unter https://zenodo.org/doi/10.5281/zenodo.13174950.

Fox, T. F. (1965) *Crisis in communication: the functions and future of medical journals*, London, Athlone Press.

Franceschet, M.; Costantini, A. (2011) „The first Italian research assessment exercise: A bibliometric perspective", *Journal of Informetrics*, Vol. 5, No. 2, S. 275–291 [Online]. DOI: https://doi.org/10.1016/j.joi.2010.12.002.

Fraumann, G.; Mutz, R. (2022) „3.4 The h-index", in Ball, R. (Hg.) *Handbook bibliometrics* [Online], Berlin, Boston, De Gruyter Saur, S. 169–178. Verfügbar unter https://doi.org/10.1515/9783110646610.

Gallo, S. A.; Glisson, S. R. (2018) „External Tests of Peer Review Validity Via Impact Measures", *Frontiers in Research Metrics and Analytics*, Vol. 3 [Online]. DOI: https://doi.org/10.3389/frma.2018.00022.

Godlee, F. (2002) „Making reviewers visible: openness, accountability, and credit", *JAMA*, Vol. 287, No. 21, S. 2762–2765 [Online]. DOI: https://doi.org/10.1001/jama.287.21.2762.

Goldfinch, S.; Yamamoto, K. (2012) *Prometheus assessed?: Research measurement, peer review, and citation analysis*, Oxford, Chandos Pub.

Gorraiz, J. (2009) *Bibliometrie* [Online]. Verfügbar unter https://homepage.univie.ac.at/juan.gorraiz/Citation/bibliometrie.htm (Abgerufen am 15 August 2024).

Gould, T. H. P. (2013) *Do we still need peer review? An argument for change*, Lanham, Md, Scarecrow Press.

Gross, P. L.; Gross, E. M. (1927) „College Libraries and Chemical Education", *Science*, Vol. 66, No. 1713, S. 385–389 [Online]. DOI: https://doi.org/10.1126/science.66.1713.385.

Haustein, S. (2016) „Grand challenges in altmetrics: heterogeneity, data quality and dependencies", *Scientometrics*, Vol. 108, No. 1, S. 413–423 [Online]. DOI: https://doi.org/10.1007/s11192-016-1910-9.

Haustein, S.; Larivière, V. (2015) „The Use of Bibliometrics for Assessing Research: Possibilities, Limitations and Adverse Effects", in Welpe, I.

M., Wollersheim, J., Ringelhan, S. & Osterloh, M. (Hg.) *Incentives and Performance,* Cham, Springer International Publishing, S. 121–139.

Hirsch, J. E. (2005) „An index to quantify an individual's scientific research output", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 102, No. 46, S. 16569–16572 [Online]. DOI: https://doi.org/10.1073/pnas.0507655102.

Hornbostel, S.; Böhmer, S.; Klingsporn, B.; Neufeld, J.; Ins, M. von (2009) „Funding of young scientist and scientific excellence", *Scientometrics*, Vol. 79, No. 1, S. 171–190 [Online]. DOI: https://doi.org/10.1007/s11192-009-0411-5.

Jennings, C. (2006) „Quality and value: The true purpose of peer review", *Nature* [Online]. DOI: https://doi.org/10.1038/nature05032.

Jensen, P.; Rouquier, J.-B.; Croissant, Y. (2009) „Testing bibliometric indicators by their prediction of scientists promotions", *Scientometrics*, Vol. 78, No. 3, S. 467–479 [Online]. DOI: https://doi.org/10.1007/s11192-007-2014-3.

Krattenthaler, C. (2021) „Was der h-Index wirklich aussagt", *Mitteilungen der Deutschen Mathematiker-Vereinigung*, Vol. 29, No. 3, S. 124–128 [Online]. DOI: https://doi.org/10.1515/dmvm-2021-0050.

Kulczycki, E. (2023) *The evaluation game : how publication metrics shape scholarly communication*, Cambridge University Press.

Lee, C. J.; Sugimoto, C. R.; Zhang, G.; Cronin, B. (2013) „Bias in peer review", *Journal of the American Society for Information Science and Technology*, Vol. 64, No. 1, S. 2–17 [Online]. DOI: https://doi.org/10.1002/asi.22784.

Lotka, A. J. (1926) „The frequency distribution of scientific productivity", *Journal of the Washington Academy of Sciences*, Vol. 16, No. 12, S. 317–323 [Online]. Verfügbar unter http://www.jstor.org/stable/24529203.

MacRoberts, M. H.; MacRoberts, B. R. (1986) „Quantitative Measures of Communication in Science: A Study of the Formal Level", *Social Studies of Science*, Vol. 16, No. 1, S. 151–172 [Online]. DOI: https://doi.org/10.1177/030631286016001008.

Mahdi, S.; D'Este, P.; Neely, A. (2008) „Citation Counts: Are They Good Predictors of RAE Scores? A bibliometric analysis of RAE 2001". London: AIM Report.

Marx, W.; Bornmann, L. (2015) „Bibliometrische Verfahren zur Bewertung von Forschungsleistung", *Soziale Welt*, Vol. 66, No. 2, S. 161–176 [Online]. DOI: https://doi.org/10.5771/0038-6073-2015-2-161.

McCook, A. (2016) *Dear peer reviewer, you stole my paper: An author's worst nightmare* [Online]. Verfügbar unter https://retractionwatch.com/2016/12/12/dear-peer-reviewer-stole-paper-authors-worst-nightmare/ (Abgerufen am 20 August 2024).

McDonald, R. J.; Cloft, H. J.; Kallmes, D. F. (2007) „Fate of submitted manuscripts rejected from the American Journal of Neuroradiology: outcomes and commentary", *American journal of neuroradiology*, Vol. 28, No. 8, S. 1430–1434 [Online]. DOI: https://doi.org/10.3174/ajnr.A0766.

McDonald, R. J.; Cloft, H. J.; Kallmes, D. F. (2009) „Fate of manuscripts previously rejected by the American Journal of Neuroradiology: a follow-up analysis", *American journal of neuroradiology*, Vol. 30, No. 2, S. 253–256 [Online]. DOI: 10.3174/ajnr.A1366.

McKiernan, E. C.; Schimanski, L. A.; Muñoz Nieves; C., Matthias, L.; Niles, M. T.; Alperin, J. P. (2019) „Use of the Journal Impact Factor in academic review, promotion, and tenure evaluations", *PeerJ Preprints*, Vol. 7 [Online]. DOI: https://doi.org/10.7287/peerj.preprints.27638v2.

Melin, G.; Danell, R. (2006) „The top eight percent: development of approved and rejected applicants for a prestigious grant in Sweden", *Science and Public Policy*, Vol. 33, No. 10, S. 702–712 [Online]. DOI: https://doi.org/10.3152/147154306781778579.

Mittermaier, B. (2022) „2.1 Peer Review and Bibliometrics", in Ball, R. (Hg.) *Handbook bibliometrics* [Online], Berlin, Boston, De Gruyter Saur, S. 77–90. Verfügbar unter https://doi.org/10.1515/9783110646610.

Mittermaier, B. (2023) *Peer Review, Bibliometrie, Altmetrics – gibt es einen Königsweg bei Evaluierungen?* [Online]. Verfügbar unter https://doi.org/10.34734/FZJ-2023-03664.

Moed, H. F. (2016) „Altmetrics as Traces of the Computerization of the Research Process", in Sugimoto, C. R. (Hg.) *Theories of Informetrics and Scholarly Communication,* De Gruyter Saur, S. 360–371.

Moed, H. F.; Bruin, R. E. de; van Leeuwen, T. N. (1995) „New bibliometric tools for the assessment of national research performance: Database description, overview of indicators and first applications", *Scientometrics*, Vol. 33, No. 3, S. 381–422 [Online]. DOI: https://doi.org/10.1007/BF02017338.

Nederhof, A. J.; van Raan, A. (1993) „A bibliometric analysis of six economics research groups: A comparison with peer review", *Research Policy*, Vol. 22, No. 4, S. 353–368 [Online]. DOI: https://doi.org/10.1016/0048-7333(93)90005-3.

Norris, M.; Oppenheim, C. (2003) „Citation counts and the Research Assessment Exercise V", *Journal of Documentation*, Vol. 59, No. 6, S. 709–730 [Online]. DOI: https://doi.org/10.1108/00220410310698734.

Ochsner, M. (2022) „2.6 Relationship between Peer Review and Bibliometrics", in Ball, R. (Hg.) *Handbook bibliometrics* [Online], Berlin, Boston, De Gruyter Saur, S. 125–132. Verfügbar unter https://doi.org/10.1515/9783110646610.

Oppenheim, C. (1997) „The correlation between citation counts and the 1992 research assessment exercise ratings for British research in genetics, anatomy and archaeology", *Journal of Documentation*, Vol. 53, No. 5, S. 477–487 [Online]. DOI: https://doi.org/10.1108/EUM0000000007207.

Opthof, T.; Furstner, F.; van Geer, M.; Coronel, R. (2000) „Regrets or no regrets? No regrets! The fate of rejected manuscripts", *Cardiovascular research*, Vol. 45, No. 1, S. 255–258 [Online]. DOI: https://doi.org/10.1016/s0008-6363(99)00339-9.

Paltridge, B. (2017) *The Discourse of Peer Review: Reviewing Submissions to Academic Journals*, London, Palgrave Macmillan UK.

Pölönen, J.; Engels, T.; Guns, R. (2019) „Ambiguity in identification of peer-reviewed publications in the Finnish and Flemish performance-based research funding systems", *Science and Public Policy*, Vol. 47, No. 1, S. 1–15 [Online]. DOI: https://doi.org/10.1093/scipol/scz041.

Priem, J. (2014) „Altmetrics", in Cronin, B. & Sugimoto, C. R. (Hg.) *Beyond bibliometrics: Harnessing multidimensional indicators of scholarly impact,* Cambridge, Massachusetts, MIT Press, S. 263–287.

Priem, J.; Groth, P.; Taraborelli, D. (2012) „The altmetrics collection", *PloS one*, Vol. 7, No. 11, e48753 [Online]. DOI: https://doi.org/10.1371/journal.pone.0048753.

Priem, J.; Taraborelli, D.; Groth, P.; Neylon, C. (2010) *Altmetrics: A manifesto* [Online]. Verfügbar unter https://zenodo.org/doi/10.5281/zenodo.12684248.

Pritchard, A. (1969) „Statistical bibliography or bibliometrics?", *Journal of Documentation*, Vol. 25, No. 4, S. 348–349.

Reinhart, M.; Schendzielorz, C. (2024) „Peer-review procedures as practice, decision, and governance—the road to theories of peer review", *Science and Public Policy*, Vol. 51, No. 3, S. 543–552 [Online]. DOI: https://doi.org/10.1093/scipol/scad089 (Abgerufen am 14 August 2024).

Rinia, E.; van Leeuwen, T.; van Vuren, H.; van Raan, A. (1998) „Comparative analysis of a set of bibliometric indicators and central peer review criteria", *Research Policy*, Vol. 27, No. 1, S. 95–107 [Online]. DOI: https://doi.org/10.1016/S0048-7333(98)00026-2.

Roy, R. (1985) „Funding Science: The Real Defects of Peer Review and An Alternative To It", *Science, Technology, & Human Values*, Vol. 10, No. 3, S. 73–81 [Online]. DOI: https://doi.org/10.1177/016224398501000309.

Rühle, A. (2009) „Wie misst man Exzellenz?", *Süddeutsche Zeitung*, 29. Januar [Online]. Verfügbar unter https://web.archive.org/web/20090131032736/http://www.sueddeutsche.de:80/jobkarriere/759/456427/text.

Simsek, M.; Vaan, M. de; van de Rijt, A. (2024) „Do grant proposal texts matter for funding decisions? A field experiment", *Scientometrics*, Vol. 129, No. 5, S. 2521–2532 [Online]. DOI: https://doi.org/10.1007/s11192-024-04968-7.

Smith, R. (2006) „Peer review: a flawed process at the heart of science and journals", *Journal of the Royal Society of Medicine*, Vol. 99, No. 4, S. 178–182 [Online]. DOI: https://doi.org/10.1258/jrsm.99.4.178.

Stephen, D.; Stahlschmidt, S. (2024) „Contrasting cross-correlation: Meta-analyses of the associations between citations and 13 altmetrics, incorporating moderating variables", *Scientometrics* [Online]. DOI: https://doi.org/10.1007/s11192-024-05006-2.

Sugimoto, C. R.; Work, S.; Larivière, V.; Haustein, S. (2017) „Scholarly use of social media and altmetrics: A review of the literature", *Journal of the Association for Information Science and Technology*, Vol. 68, No. 9, S. 2037–2062 [Online]. DOI: https://doi.org/10.1002/asi.23833.

Tennant, J. P., Dugan, J. M., Graziotin, D., Jacques, D. C., Waldner, F., Mietchen, D., Elkhatib, Y., B Collister, L., Pikas, C. K., Crick, T., Masuzzo, P., Caravaggi, A., Berg, D. R., Niemeyer, K. E., Ross-Hellauer, T., Mannheimer, S., Rigling, L., Katz, D. S., Greshake Tzovaras, B., Pacheco-Mendoza, J., Fatima, N., Poblet, M., Isaakidis, M., Irawan, D. E., Renaut, S., Madan, C. R., Matthias, L., Nørgaard Kjær, J., O'Donnell, D. P., Neylon, C., Kearns, S., Selvaraju, M.; Colomb, J. (2017) „A multi-disciplinary perspective on emergent and future innovations in peer review", *F1000Research*, Vol. 6, S. 1151 [Online]. DOI: https://doi.org/10.12688/f1000research.12037.3.

Thelwall, M. (2021) „Measuring societal impacts of research with altmetrics? Common problems and mistakes", *Journal of Economic*

Surveys, Vol. 35, No. 5, S. 1302–1314 [Online]. DOI: https://doi.org/10.1111/joes.12381.

Thelwall, M., Kousha, K., Abdoli, M., Stuart, E., Makita, M., Wilson, P.; Levitt, J. (2023) „Do altmetric scores reflect article quality? Evidence from the UK Research Excellence Framework 2021", *Journal of the Association for Information Science and Technology*, Vol. 74, No. 5, S. 582–593 [Online]. DOI: https://doi.org/10.1002/asi.24751.

Thelwall, M., Kousha, K., Wouters, P., Waltman, L., Rijcke, S. de, Rushforth, A.; Franssen, T. (2015) *The metric tide: Literature review* [Online]. Verfügbar unter https://doi.org/10.13140/RG.2.1.5066.3520.

Traag, V. A.; Waltman, L. (2019) „Systematic analysis of agreement between metrics and peer review in the UK REF", *Palgrave Communications*, Vol. 5, No. 1 [Online]. DOI: https://doi.org/10.1057/s41599-019-0233-x.

Tunger, D. (2022) „3.3 The Journal Impact Factor: A Bibliometric Indicator with a Long Past", in Ball, R. (Hg.) *Handbook bibliometrics* [Online], Berlin, Boston, De Gruyter Saur, S. 159–168. Verfügbar unter https://doi.org/10.1515/9783110646610.

Tunger, D.; Ahn, H.; Clermont, M.; Krolak, J.; Langner, J.; Mittermaier, B. (2022) „Altmetrics as Additional Indicators for Analyzing Research Efficiency of Universities? : A Conceptual and Empirical Study", *Betriebswirtschaftliche Forschung und Praxis*, Vol. 74, No. 3, 262-280 [Online]. DOI: https://doi.org/10.34734/FZJ-2022-05904.

Tunger, D.; Ahn, H.; Clermont, M.; Krolak, J.; Meier, A. (2022) „3.2 From Simple Publication Figures to Complex Indicators: 3.2 Bibliometrics and the Dilemma of Methodological Correctness, Significance, and Economic Necessity", in Ball, R. (Hg.) *Handbook bibliometrics* [Online], Berlin, Boston, De Gruyter Saur. Verfügbar unter https://doi.org/10.1515/9783110646610.

Tunger, D.; Meier, A. (2018) „Altmetrics: Ein innovativer Service für Bibliotheken", in Schade, F. & Georgy, U. (Hg.) *Praxishandbuch Informationsmarketing,* De Gruyter, S. 292–302.

Tunger, D.; Meier, A. (2022) „4.1 The Future Has Already Begun: Origin, Classification, and Applications of Altmetrics in Scholarly Communication", in Ball, R. (Hg.) *Handbook bibliometrics* [Online], Berlin, Boston, De Gruyter Saur, S. 181–190. Verfügbar unter https://doi.org/10.1515/9783110646610.

van Besselaar, P.; Leydesdorff, L. (2009) *Past Performance as Predictor of Successful Grant Applications: A Case Study* [Online]. Verfügbar unter https://doi.org/10.48550/arXiv.0911.3085.

van Leeuwen, T. N.; Moed, H. F. (2012) „Funding decisions, peer review, and scientific excellence in physical sciences, chemistry, and geosciences", *Research Evaluation*, Vol. 21, No. 3, S. 189–198 [Online]. DOI: https://doi.org/10.1093/reseval/rvs009.

van Raan, A. F. J. (2005) „Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods", *Scientometrics*, Vol. 62, No. 1, S. 133–143 [Online]. DOI: https://doi.org/10.1007/s11192-005-0008-6.

van Raan, A. F. J. (2006) „Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups", *Scientometrics*, Vol. 67, No. 3, S. 491–502 [Online]. DOI: https://doi.org/10.1556/Scient.67.2006.3.10.

Vickery, B. C. (1948) „Bradford's Law of Scattering", *Journal of Documentation*, Vol. 4, No. 3, S. 198–203 [Online]. DOI: https://doi.org/10.1108/eb026133.

Vieira, E. S.; Cabral, J. A.; Gomes, J. A. (2014) „Definition of a model based on bibliometric indicators for assessing applicants to academic positions", *Journal of the Association for Information Science and Technology*, Vol. 65, No. 3, S. 560–577 [Online]. DOI: https://doi.org/10.1002/asi.22981.

Vieira, E. S.; Gomes, J. A. N. F. (2018) „The peer-review process: The most valued dimensions according to the researcher's scientific career", *Research Evaluation*, Vol. 27, No. 3, S. 246–261 [Online]. DOI: https://doi.org/10.1093/reseval/rvy009.

Walsh, E.; Rooney, M.; Appleby, L.; Wilkinson, G. (2000) „Open peer review: a randomised controlled trial", *The British journal of psychiatry: the journal of mental science*, Vol. 176, S. 47–51 [Online]. DOI: https://doi.org/10.1192/bjp.176.1.47.

Warner, J. (2000) „A critical review of the application of citation studies to the Research Assessment Exercises", *Journal of Information Science*, Vol. 26, No. 6, S. 453–459 [Online]. DOI: https://doi.org/10.1177/016555150002600607.

Weller, A. C. (2001) *Editorial peer review: Its strengths and weaknesses*, Medford, N.J., Information Today.

Wennerås, C.; Wold, A. (1997) „Nepotism and sexism in peer-review", *Nature*, Vol. 387, No. 6631, S. 341–343 [Online]. DOI: https://doi.org/10.1038/387341a0.

KAPITEL 7

Wenn die Auswahl der Experten Expertise erfordert: Die Debatte um den Ursprung von SARS-CoV-2 und das pervasive Verhältnis von Vertrauen und Wissen

Andreas Kaminski und Jörn Wiengarn

1. Einleitung

Sachbezogener Dissens und Kontroversen gehören zum selbstverständlichen und vitalen wissenschaftlichen Alltag. Zumeist bleiben sie dabei eine wissenschafts*interne* Angelegenheit: ausgetragen von Wissenschaftlerinnen in wissenschaftlichen Zirkeln, Zeitschriften und Journalen. Dissens erfüllt eine wissenschaftliche Funktion, indem er der Überprüfung von Hypothesen und Theorien dienen, das Verständnis verfeinern sowie den Denk- und

Zitationsvorschlag für dieses Kapitel:
Kaminski, A. und Wiengarn, J. 2026. Wenn die Auswahl der Experten Expertise erfordert: Die Debatte um den Ursprung von SARS-CoV-2 und das pervasive Verhältnis von Vertrauen und Wissen. In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* S. 169–221. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.i. License: CC BY 4.0

Lösungsraum erweitern kann. Aus einer wissenschafts*externen* Perspektive werden wissenschaftliche Debatten häufig erst dann wahrgenommen, wenn wissenschaftliche Expertise eine Funktion in der Orientierungs- und Entscheidungsfindung von Personen, Organisationen, Systemen oder der Öffentlichkeit spielen soll. In einer solchen Situation jedoch können wissenschafts*interne* Kontroversen – wissenschafts*extern* wahrgenommen – das Vertrauen in die Erfüllung der gesellschaftlichen Funktion irritieren oder mindern.

Dissens und Kontroversen in den Wissenschaften setzen einen inhaltsbezogenen Streit zwischen Fachvertreterinnen voraus. In der Öffentlichkeit ruft dies die Fragen auf: Wie vertrauenswürdig sind die jeweiligen Wissenschaftler? Handelt es sich jeweils um Expertinnen? Wer hat die größere Expertise? Und wie sind daher die Kontroversen einzuschätzen?

Solche Fragen sind auch deshalb so drängend, da vieles, das meiste sogar, das wir wissen oder zu wissen glauben, wir nur deshalb zu wissen behaupten können, da es uns andere mitgeteilt haben – und nicht selten eben solche, die wir zumindest für Experten halten. Damit betreffen solche Fragen das Verhältnis von Vertrauen und Wissen. John Hardwig hat dieses Verhältnis auf die Formel gebracht, dass Vertrauen Wissen stützt und unterstützt: „We do not normally notice the air we breathe. Similarly, we epistemologists have not noticed the climate of trust that is required […] to support much of our knowledge." (Hardwig 1991: 693) Doch welches Verhältnis zwischen Vertrauen und Wissen besteht, ist damit noch nicht geklärt. Im Gegenteil: Hardwigs Überlegung mag zugestimmt werden, und sie kann doch zu ganz unterschiedlichen Ideen über das Verhältnis von Wissen und Vertrauen führen.

Daher wollen wir diese beiden Fragen aufwerfen:

1. Wie lässt sich das Verhältnis von Vertrauen und Wissen begreifen?
2. Und wie können im Angesicht eines Dissenses zwischen Expertinnen Laien entscheiden, wem sie vertrauen und so Glauben schenken sollen?

Wir gehen auf zweierlei Weise vor: Wir werden diese Fragen einerseits begrifflich untersuchen, uns andererseits aber auch intensiv mit einem brisanten und reichhaltigen Fallbeispiel beschäftigen – dem Expertenstreit über den Ursprung von SARS-CoV-2. Hier findet sich ein Expertendissens in der Frage, ob das Virus aus einem Labor stammt oder nicht. Dazu gleich mehr.

Wie kann rational entschieden werden, welchen Expertinnen vertraut werden sollte? In der sozialen Epistemologie und angrenzend daran der Philosophie von Vertrauen ist ein Vorschlag entwickelt worden, wie die Bewertung und Auswahl der Expertinnen durch Laien rational vorgenommen werden kann. Dieser basiert auf der Idee, dass Laien durch höherstufige *Evidenz* darüber informiert werden können, wer vertrauenswürdig(er) ist. Der Grundgedanke gleicht *einem* Lösungsansatz innerhalb der Philosophie von Vertrauen[1]: Die epistemischen Theorien von Vertrauen betrachten das *Vertrauensproblem*, nämlich wem vernünftigerweise vertraut werden kann, als ein *Wissensproblem*. Im Kontext einer Bezeugung – wenn also A gegenüber B behauptet, dass p – stellt sich die Frage, ob B rationalerweise A vertrauen kann. Zwar kann man im Kontext einer solchen Bezeugung in der Regel nicht direkt verifizieren, ob das Behauptete, also p, auch wahr ist – sonst würde es sich ja auch nicht um ein *Vertrauens*problem handeln. Aber es gibt – so die Grundannahme der epistemischen Ansätze – ein höherstufiges Wissen, es gibt Evidenz 2. Ordnung, nämlich über A, die Redesituation oder weitere Umstände, die Evidenz für die Glaubwürdigkeit von A oder p bieten kann. Dabei wird folglich angenommen, dass das Vertrauensproblem durch höherstufiges Wissen – und damit *epistemisch* – gelöst werden kann.

In den letzten 30 Jahren sind jedoch grundlegende Einwände gegen eine rein epistemische Lösung des Vertrauensproblems entwickelt worden.[2] Belasten diese auch die epistemischen Versuche, die Expertise der Experten zu bewerten? Dieser Frage gehen wir im Folgenden nach, indem wir uns zunächst (Abschn. 2) mit dem klassischen Ansatz von Alvin I. Goldman, wie Laien die Vertrauenswürdigkeit von Expertinnen epistemisch bewerten können, befassen. Anschließend (Abschn. 3) gehen wir auf die Einwände

gegen die epistemische Lösung des Vertrauensproblems ein. Wir werden diese Einwände gleich in einer auf das Problem passenden Weise formulieren: Nämlich als Antworten auf die Frage, wie das Verhältnis von Wissen (Evidenz) und Vertrauen begriffen werden kann. Wir wollen zeigen, dass die Versuche, eine klare Demarkationslinie zwischen Vertrauen und Wissen zu ziehen, vor einige Schwierigkeiten gestellt sind. Angesichts dieser führen wir eine alternative Konzeption ihres Verhältnisses vor: Vertrauen und Wissen durchdringen einander. Dieses Pervasivitätsmodell bietet eine Alternative zu dem stillschweigend häufig unterstellten Demarkationsansatz.

Im anschließenden Teil (Abschn. 4) wenden wir uns einem exemplarischen Fall einer wissenschaftlichen Kontroverse zu: der Frage nach dem Ursprung von SARS-CoV-2. In dieser Debatte werden zwei unterschiedliche Hypothesen über die Herkunft des Coronavirus (Labor vs. Natur) vertreten. In der Öffentlichkeit ist dieser Fall mit viel Aufmerksamkeit verfolgt worden, weil die Antwort auf diese Frage große rechtliche, politische und wissenschaftspolitische Folgen hat. Wir werden einige weitere mögliche Kriterien diskutieren, die Nicht-Expertinnen heranziehen können, um sich in der Frage zu orientieren – Kriterien, die über den ursprünglich von Goldman vorgeschlagenen Katalog hinausgehen. Wir glauben jedoch zeigen zu können (Abschn. 5), dass auch diese Kriterien die Pervasivität von Vertrauen und Wissen nicht auflösen – also Wissen und Vertrauen nicht eindeutig und restlos voneinander trennen können. Die Fallstudie liefert damit eine Illustration dafür, inwiefern es nicht möglich scheint, das Vertrauensproblem auf ein reines Wissensproblem zurückzuführen.

2. Das „Novize-2-Experten-Problem“ & Goldmans Lösungsvorschlag

Das Problem, um das es im Folgenden geht, lässt sich, wie folgt, umreißen: Angenommen, wir haben zwei Expertinnen (im Folgenden: E_1 und E_2), deren Antworten auf dieselbe wissenschaftliche

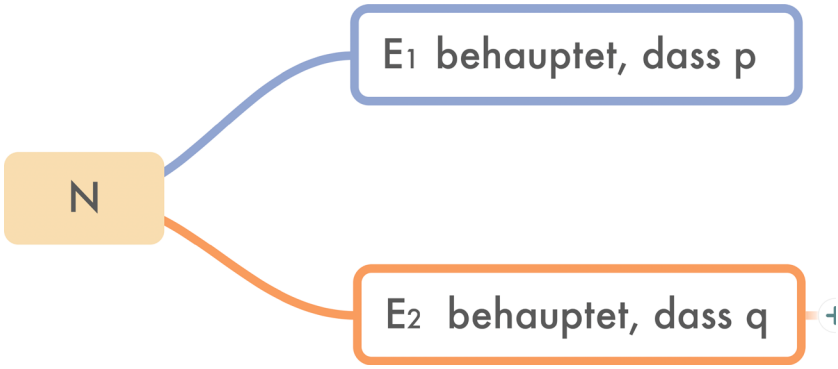


Abbildung 7.1: Das Ausgangsproblem.

Frage konträr sind oder einander widersprechen.[3] Kann ein Nicht-Experte, ein *Novize* (N) in den Worten des Philosophen Alvin I. Goldmans, in dieser Situation eine rationale Entscheidung darüber treffen, ob einer Expertin und, wenn ja, welcher vertraut werden kann?

Goldman hat eine Strategie entwickelt, die Laien in die Lage versetzen soll, eine vernünftige Antwort auf diese Frage zu finden. Dabei betont Goldman, dass sein Ansatz gewisse Idealisierungen, die in der Epistemologie üblich seien, vermeiden will. Anstelle von epistemischen Akteuren mit unbegrenzten zeitlichen oder intellektuellen Ressourcen will er realitätsnah gedachte Novizen betrachten, die sich einem wissenschaftlichen Dissens gegenübersehen und die jeweiligen Expertinnen bewerten wollen.

> „The task for the layperson who is consulting putative experts, and who hopes thereby to learn a true answer to the target question, is to decide who has superior expertise, or who has better deployed his expertise to the question at hand. The novice/2-experts problem is whether a layperson can justifiably choose one putative expert as more credible or trustworthy than the other with respect to the question at hand, and what might be the epistemic basis for such a choice?" (Goldman 2001: 92)

Ein genereller Ansatzpunkt, diese Aufgabe zu lösen, besteht darin, Wissen über die Zeuginnen zu gewinnen. Der Novize könnte also

versuchen, mehr über sie in Erfahrung zu bringen. Stellen wir etwa fest, dass eine Zeugin eine von ihr beschriebene Situation gar nicht richtig wahrnehmen konnte, weil sie eine eingeschränkte Sicht darauf hatte, dann verändert dieses Wissen den Grad der Vertrauenswürdigkeit, die die Zeugin für uns hat. „My point is, however, that your evidence about the properties of the speaker is crucial evidence for your overall entitlement to accept the speaker's assertion. […] So what you empirically determine about a speaker can make a massive difference to your overall justifiedness in accepting his utterances." (Ebd.: 88) Was mit Blick auf diese generelle, alltägliche Situation gilt, lässt sich, so Goldman, auch auf das Ausgangsproblem, den Expertinnendissens anwenden:

> „The same obviously holds about two putative experts, who make conflicting claims about a given subject-matter. Which claim you should accept (if either) can certainly be massively affected by your empirical discoveries about their respective abilities and opportunities to know the truth of the matter (and to speak sincerely about it)." (Ebd.: 88f.)

Goldman geht also davon aus, dass sich das Problem, welcher Expertin vertraut werden darf, als ein Wissensproblem darstellt, das entsprechend, wenn überhaupt, durch Wissen gelöst werden kann. Dazu entwickelt er ein Set von Kriterien mit dem Anspruch, dass diese auch für Laien anwendbar seien. Der letzte Punkt ist wichtig, da es ihm darum geht, dass die Kriterien anwendbar sein sollen, ohne dass man selbst Expertin im fraglichen Themengebiet ist. Der Novize kann sich schließlich nicht direkt – aufgrund eigener Expertise – zu den Behauptungen von E_1 und E_2 verhalten. Anders gesagt: N kann nicht aufgrund eigenen Wissens über die fragliche Domäne die Wahrheit von p oder q direkt bewerten (ebd.: 90).

Expertinnen sind für Goldman Personen, die mehr wahre Überzeugungen sowie weniger falsche als Laien in einer spezifischen Domäne haben. Sie besitzen jedoch nicht bloß mehr zutreffende Überzeugungen, sie verfügen auch über entsprechende Fähigkeiten und Fertigkeiten, neue Informationen zu analysieren und zu

bewerten – Fähigkeiten, über die Novizen nicht verfügen (ebd.: 91f.). Zudem verfügen Expertinnen nicht bloß über Wissen über die vorhandene Evidenz in der Expertendomäne, sondern auch über Wissen über Haltungen und Urteile anderer Expertinnen in diesem Feld zur vorhandenen Evidenz (ebd.: 92). All dies muss folglich bei den Novizen ausgeschlossen werden.

Goldman führt fünf mögliche Quellen von Evidenz an, die Laien in einer solchen Situation nutzen können (ohne dass er damit behaupten möchte, die Liste sei vollständig). Wir gehen seine Vorschläge in der Reihenfolge, in der er sie nennt, durch und diskutieren sogleich die Voraussetzungen (und Probleme) ihrer Anwendung durch Novizen, wie sie Goldman bereits erörtert.

1. Argumente: Expertinnen führen in der Regel Argumente an, um ihre Behauptungen zu stützen. Argumente bestehen aus Aussagen, zwischen denen eine stützende Beziehung besteht: Einige Aussagen rechtfertigen eine andere, etwa in der Art: diese Daten stützen jene Hypothese. Laien können diese Argumente zumeist nicht *direkt* bewerten; sie können ihre Gültigkeit nicht eigenständig beurteilen. Die Beurteilung, ob die Prämissen zutreffend sind und ob aus ihnen das Behauptete folgt, erfordert in der Regel ein hohes Maß an Expertise, das weit über Fragen der formalen Logik hinausgeht. Für Goldman gibt es jedoch eine Alternative. Argumente können nämlich nicht nur direkt, sondern auch indirekt bewertet werden. Die *indirekte* Bewertung erfolgt durch die Beobachtung der *Weise*, in der Argumente vorgebracht werden. In der Weise, wie eine Expertin argumentiert, mag sie einer anderen als überlegen *erscheinen* – etwa schneller, bestimmter oder ‚smoother' antworten oder dergleichen. Es handelt sich mit anderen Worten um Indikatoren, die auf Observablen beruhen. Goldman hebt selbst hervor, dass es sich dabei um eine problematische Strategie handeln mag:

> „Of course, quickness and smoothness are problematic indicators of informational mastery. Skilled debaters and well-coached witnesses can appear better-informed because of their stylistic polish, which is not a true indicator of superior expertise. This

> makes the proper use of indirect argumentative justification a very delicate matter." (Ebd.: 96)

Goldman hält dies gleichwohl für eine mögliche Bewertungsstrategie für Novizen. *Faktisch* mag dies zutreffend sein: Es gibt sicherlich viele Situationen, in denen Novizen sich so verhalten und die Argumente der Expertinnen anhand der argumentativen Performance indirekt bewerten. Aber *sollten* sie dies tun? Wir glauben, dass es sich dabei um eine ungemein täuschungsanfällige Strategie handelt, und zwar aus zwei Gründen, die z.T. über die von Goldman selbst genannten hinaus gehen:

Es mag zum einen kontextspezifische Gründe geben, warum die Argumentation einer vermeintlichen Expertin besser als die der anderen erscheint – ohne dies tatsächlich zu sein. Ein solcher Fall wäre z.B., wenn eine Expertin eine Frage gestellt bekommt, die eine längere, komplizierte Antwort erfordert. E_1 verheddert sich beim Versuch, diese Komplexität in einer Weise, dass sie für Laien nachvollziehbar ist, darzustellen. Sie setzt zwei- bis dreimal an, sie gibt zwar korrekte Antworten, aber sie bleiben entweder nicht nachvollziehbar oder sind erkennbar Stückwerk, indem sie beim Versuch der Erläuterung abbricht. E_2 mag diese Komplexität ungerechtfertigt übergehen und eine problematische Simplifikation bieten. Ihre Argumentation erscheint daher ‚smoother', ohne deshalb ‚wahr' oder besser ‚begründet' zu sein.

Zum anderen dürften in der diskursiven Performance-Erscheinung von E_1 und E_2 auch andere Züge eine Rolle spielen, die aber keine indikatorische Beziehung zur Stützung von p oder q aufweisen. Betrachten wir einen extrem wirkenden Fall, in dem E_1 arrogant gegenüber dem Argument von E_2 auftritt. Es ist aus unserer Sicht nicht klar, ob bzw. welche epistemische Beziehung zwischen der arrogant wirkenden Erscheinungsweise und der Bewertung des Arguments besteht. Vielleicht ist das Argument von E_2 für eine Person, die es einschätzen kann, in der Tat lächerlich. Vielleicht hat E_1 zu häufig schon solche Argumente gehört und kann es in diesem Moment nicht verbergen, dass sie sie für lächerlich hält. Vielleicht ist E_1 aber auch überfordert mit der Situation, und sie

agiert sozial einfach ungeschickt; vielleicht ist das Argument von E_2 dem von E_1 einfach überlegen, und E_1 weiß sich nicht anders zu helfen, als arrogant darüber hinwegzulächeln.

Kurz: Gibt es wirklich eine belastbare epistemische Relation zwischen der argumentativen Präsentationsweise der Expertinnen und dem Wahrheitswert des Arguments, die Novizinnen nutzen könnten? Es scheint aus unserer Sicht hier keine allgemeine Regel klar erkennbar zu sein.[4] Dass Laien *faktisch* die Erscheinungsweise zur Bewertung der Argumente nutzen, dürfte zutreffend sein. Es ist auch insofern verständlich, als ihnen die direkte Bewertung ja gerade nicht zugänglich ist. Dass Novizen Expertise auf diese Weise bewerten *sollten* – und dies ist Goldmans Behauptung – erscheint uns aber fraglich (Coady/Corry 2013: 24–26; Levy 2024: 2081).

2. Zustimmung anderer Experten: Eine andere Strategie besteht darin, die Vertrauenswürdigkeit von Expertinnen auf der Basis ihrer Bewertung durch andere Experten zu bewerten. Gibt es weitere Experten, die E_1 oder E_2 zustimmen? Wie viele stimmen ihnen jeweils zu? „Won't N be fully justified in trusting E_1 over E_2 if almost all other experts on the subject agree with E_1, or if even a preponderance of the other experts agree with E_1?" (Goldman 2001: 97) Wiederum erscheint dies als eine naheliegende, vernünftige Strategie. Ihre Anwendung ist jedoch keineswegs so einfach und verlässlich, wie man auf den ersten Blick meinen mag. Gegen die Verlässlichkeit dieses Kriteriums sprechen etwa historische Irrtümer, denen der wissenschaftliche Konsens aufgesessen ist. Ein Beispiel stellt die Frage homosexuellen Verhaltens unter Tieren dar. Die meisten Experten waren früher der Ansicht, dass es dies nicht gibt, obgleich es heute gut dokumentiert ist. – Goldman selbst sieht einen Teil dieser Schwierigkeiten (vgl. sein Guru-Argument oder das Problem der Gerüchte), weshalb er ein differenzierteres Kriterium einführt.

3. Meta-Experten: Goldman schlägt ergänzend vor, die Einschätzung sog. Meta-Experten zu verwenden. Meta-Experten bewerten die Expertise von Expertinnen. „N should seek evidence about the two rival experts' relative degrees of expertise by consulting third parties' assessments of their expertise. If 'meta-experts' give E_1

higher 'ratings' or 'scores' than E_2, shouldn't N rely more on E_1 than E_2?" (Ebd.: 97) Zu dieser Bewertung zählt auch die Vergabe von akademischen Titeln, Akkreditierungen, Zertifizierungen und dergleichen. Diese „credentials" geben Experten Glaubwürdigkeit.

Dieses Kriterium weist eine Reihe von Problemen auf. Eines betrifft seinen Status in Zeiten wissenschaftlicher Umbrüche (besonders dann, wenn man noch nicht weiß, ob es sich um einen Umbruch handeln wird oder nicht) oder auch in Frühphasen einer Wissenschaft. Ein Beispiel hierfür wäre die Kontinentaldrifttheorie, die zu Beginn überwiegend auf Ablehnung stieß (und nicht bloß aus nichtepistemischen Gründen). Wissenschaftler, die eine große Reputation besaßen, lehnten die These, dass die Kontinente sich horizontal bewegten, ab. Gleichwohl setzte sich diese Theorie mit der Zeit durch und ist heute unumstritten.

Das allgemeine Problem dieses Kriteriums ist jedoch seine offensichtliche Zirkularität. Sie wird sichtbar an der offenkundigen Frage, wer die Meta-Meta-Experten sind, die die Expertise der Meta-Experten bewerten. Die Zirkularität besteht darin, dass die Bewertung der Expertinnen Expertise voraussetzt.

Hier kommen wir zum Kernproblem von Goldmans Ansatz (vgl. Levy 2024)[5]: Er möchte Vertrauen auf der Basis von Wissen begründen – und das heißt: er möchte vernünftiges Vertrauen auf Wissen zurückführen. Dies ist aus unserer Sicht nicht möglich. Vertrauen und Wissen durchdringen einander intimer, als es die Idee einer klaren Schnittstelle, an der Vertrauen endet und Wissen beginnt, nahelegt (dazu mehr im nächsten Abschnitt). Das zeigt sich auch mit Blick auf die beiden weiteren Kriterien Goldmans: Evidenz über Interessenskonflikte und Track-Records.

4. Evidenz für Interessenskonflikte: Evidenz für Interessenskonflikte oder den Bias von Wissenschaftlern soll wiederum das Misstrauen in Expertinnen durch Wissen begründen. Erneut zeigt sich das gleiche Problem: Evidenz über Interessenskonflikte oder Voreingenommenheit (Bias) bewerten zu können setzt Expertise voraus. Forschung wird in aller Regel gefördert, sei es staatlich, sei es durch gemeinnützige Einrichtungen, sei es durch privatwirtschaftliche Unternehmen. Die Förderung durch privatwirtschaftliche Unternehmen stellt per se keinen Grund dar, Forschungsergebnisse

abzulehnen. Es ist kein generelles Ausschlusskriterium, aber das Wissen darum ist deshalb auch nicht irrelevant. Es kommt hier auf eine genauere Analyse an. Diese wird in der Regel wiederum nicht einfach eigenständig erfolgen können, sondern setzt Expertise voraus und genaue Kenntnisse des jeweiligen Sachverhaltes – und damit Expertinnen (Levy 2024: 2086).

5. Track-Records: Dieselbe Kritik lässt sich auch mit Bezug auf Track-Records (d.h. die Feststellung der bisherigen ‚Erfolgsgeschichte' der Expertinnen) formulieren. Ein Beispiel hierfür sind Prognosen von Experten. Es ist sinnvoll, die Urteile von Experten über Covid-19 zu bewerten, wie es etwa in einem journalistischen Beitrag über Hendrik Streeck, der während der Pandemie sehr präsent in den Medien war, unternommen wird (vgl. Schwägerl/Budde 2021). Dabei wird in dem Artikel eine Art Erfolgsanalyse (Track-Record) über Einschätzungen (Urteile, Empfehlungen, Prognosen) von Streeck erstellt. In elf wichtigen Fällen[6] soll gezeigt werden, dass sich dieser geirrt habe. Diskutiert wird unter anderem die Tödlichkeit von Covid-19, von der es heißt, dass Streeck sie massiv unterschätzt habe. Unser Punkt ist nicht zu bezweifeln, dass Streeck sich (auch) in diesem Punkt irrte. Noch geht es uns darum zu behaupten, dass sich angesichts der vielen Messprobleme (etwa die unterschiedlichen Zählweisen) gar nichts über die Sterblichkeitsrate der mit dem Virus Infizierten sagen ließe. Die sog. *Übersterblichkeit* gibt einen Weg an, die Tödlichkeit von Covid-19 einzuschätzen, auch sie muss aber im Zusammenhang mit anderen Informationen und Messverfahren beurteilt werden (Veränderungen der Demographie, andere mögliche Todesursachen). Track-Records spielen hier eine wichtige Rolle, aber sie bieten keine von Vertrauen unabhängige Basis zur Bewertung der Vertrauenswürdigkeit. Vielmehr setzen sie Vertrauen in andere voraus, die diese Track-Records erstellt und überprüft haben. Erneut – und kaum überraschend – wird damit Expertise für die Bewertung von Expertise benötigt.

Das zeigt sich auch mit Blick auf einen letzten Punkt: Nicht alle falschen Prognosen sprechen gleichermaßen für die mangelnde Expertise einer Person. Prognosen schließen zumeist andere Verläufe nicht aus. Das gilt bspw. auch für Wettervorhersagen. Es gilt nicht minder für Prognosen unter großer Unsicherheit, wie sie

gerade zu Beginn der Pandemie bestand, die aus Handlungsdruck jedoch dennoch gefordert werden. Dies bedeutet nicht, dass alle Prognosen gleichermaßen zulässig und vernünftig sind. Die Beurteilung, wann eine verfehlte Prognose an der Expertise einer Person Zweifel aufkommen lassen sollte, erfordert jedoch selbst wiederum Expertise: Man muss einschätzen können, ob es sich um einen vorwerfbaren Fehler handelte oder nur um einen nie auszuschließenden alternativen Verlauf; wenn es sich um einen Fehler handelte, um welche Art von Fehler usw.

Die verschiedenen Kritikpunkte an Goldmans Kriterien teilen eine Gemeinsamkeit: Häufig setzt die Anwendung dieser Kriterien ein Maß an Expertise voraus, das Novizen nicht erfüllen. Sie können die Kriterien daher nur dadurch vernünftig anwenden, dass sie diese Expertise von außen einbeziehen – also von anderen Experten: Um zu entscheiden, wie gut der Track-Record einer Expertin ist, wie viele Expertinnen einem Experten zustimmen, ob Interessenskonflikte bestehen, die problematisch sind – in all diesen Fällen muss der Novize der Einschätzung anderer Expertinnen vertrauen. Ausdrücklich kommt dies bei den Meta-Experten zum Vorschein. Dadurch entsteht aber eine Zirkelstruktur: N vertraut E_n in der Frage, ob man E_1 oder E_2 vertrauen sollte. Wenn nun aber N feststellt, dass es neben E_n auch E_m gibt und E_n sowie E_m zu unterschiedlichen Ergebnissen kommen, ob man E_1 oder E_2 vertrauen sollte, dann bricht dieser hintergründige Zirkel in die Verwickeltheiten der Urteilspraxis ein. Offensichtlich wird daran, dass die Reduktion des Vertrauensproblems auf ein Wissensproblem scheitert. Dies ruft die Frage auf: Wie lässt sich das Verhältnis von Wissen und Vertrauen begreifen?

3. Vertrauen und Wissen: Zwei Modelle ihres Verhältnisses

Der eben hervorgetretene Zirkel ruft den Verdacht auf, dass etwas an der Weise, wie Vertrauen und Wissen in den epistemischen

Theorien ins Verhältnis zueinander gesetzt werden, problematisch ist. Genau genommen, ist das Problem nicht auf die epistemische Theorie von Vertrauen beschränkt; diese verlängert nur die Traditionslinie einer schon mehrfach in Schwierigkeiten geratenen Fundierungstheorie des Wissens selbst. Gemäß dieser muss es ein *Fundament* des Wissens geben, das als solches eine Unabhängigkeit von anderen Erkenntnisschichten aufweisen muss, die es ja fundieren soll.

Die Fundierungstheorien des Wissens sind – lange vor der modernen Auseinandersetzung mit Vertrauen – in tiefgreifende Schwierigkeiten verwickelt worden. Die Frage etwa, was das Fundament selbst begründet, sofern es ja Wissen (und nicht ein Dogma) sein soll, deutet auf diese Schwierigkeiten hin. Wilfried Sellars hatte sie unter dem Stichwort eines Mythos des (unmittelbar) Gegebenen kritisiert (vgl. Sellars 1963).

Die epistemischen Theorien des Vertrauens weisen, wie wir gleich genauer aufzeigen werden, die gleiche Basisstruktur wie die Fundierungstheorien des Wissens auf, als sie auch davon ausgehen, dass es eine grundlegende Wissensschicht gibt, die Vertrauen fundieren kann und dabei selbst vertrauensunabhängig ist. Wie wir ebenso genauer zeigen werden, ergeben sich folgerichtig für diese ähnliche Probleme wie für die Fundierungstheorie des Wissens. Nicht ganz überraschend folgen im Übrigen auch prominente Lösungsvorschläge für das Fundierungsproblem des Wissens einer ähnlichen Grundstoßrichtung wie Theorieantworten auf das Problem der wissensfundierten Vertrauenstheorien: So entwickelt nicht nur Sellars im Ausgang der Schwierigkeit der Wissensfundierung eine holistisch-kohärentistische Antwortstrategie; ebenso schlägt auch C.A.J. Coady (1992) für die Probleme der epistemischen Vertrauenstheorien vor, dass sich Vertrauen nur im Kontext holistischer Begründungssysteme stützen kann – auf diese Lösungsvorschläge können wir aber nicht genauer eingehen. Im Folgenden fokussieren wir uns auf eine Darstellung des Fundierungsmodells des Vertrauens und seiner Probleme.

3.1. Das Fundierungs- und Demarkationsmodell

Das Fundierungsmodell nimmt an, dass Wissen unabhängig von Vertrauen ist. Dazu muss es etwas geben, was wir aus *eigener Kraft* wissen können – unabhängig von anderen. In der neuzeitlichen Philosophie waren es zwei Vorschläge, die diesbezüglich gemacht wurden: eigene Erfahrung sowie eigene Vernunfteinsicht, etwa durch logisches Schließen. Sie stellten das Fundament dar, von dem ausgehend höherstufige Schichten (zum Beispiel theoretische Aussagen) begründet werden sollten. Wenn ich etwas selbst erfahren habe, dann *weiß* ich es. Gleichermaßen weiß ich etwas, wenn ich es logisch erschlossen habe. Ich bin in diesen Fällen von anderen nicht abhängig, ich benötige mit Blick auf dieses Wissen daher auch kein Vertrauen in andere. Daher begründe und rechtfertige ich auch, wenn ich danach gefragt werde, mein behauptetes Wissen, indem ich auf *meine* Erfahrung oder auf *meine* logischen Überlegungen verweise. Ich verweise nicht auf andere. Erst jenseits dessen, was ich qua eigener Erfahrung oder eigenem logischen Nachdenken weiß oder wissen kann, beginnt der Bereich der Überzeugungen, die ich in Abhängigkeit von anderen erworben habe – anderen, denen ich vertraue. Wenn ich anderen vertraue, kann ich durch sie etwas wissen, was ich ohne sie nicht wissen würde oder gar niemals wissen könnte. Damit werden die Bereiche von Vertrauen und Wissen auf folgende Weise separiert, es gibt:

- *reines* oder *autarkes* Wissen, nämlich Wissen aus *eigener* Kraft, das auf eigener Erfahrung und eigenem logischem Nachdenken beruht; es stellt das Fundament für alles andere dar (etwa auch für theoretische Aussagen).
- *vertrauensbasiertes* Wissen, nämlich Wissen dank *anderer*, das auf meinem Vertrauen in diese anderen beruht.

Diesem ersten Fundierungsmodell liegt daher eine klare *Demarkationslinie* zugrunde. Sie trennt das reine, autarke Wissen vom vertrauensfundierten Wissen. Ihre Grenze verläuft dort, wo meine Erfahrungen und mein logisches Nachdenken enden und meine Abhängigkeit von anderen beginnt. Eine solche

Demarkationslinie scheint auch gut zu den häufig in der Sozialen Epistemologie diskutierten Fällen zu passen. Nehmen wir etwa die Frage nach meinem Wissen um einen Wahlausgang: Ich weiß nicht aus eigener Kraft, wer die letzten Wahlen gewonnen hat, da ich keine eigene Erfahrung davon habe. Ich habe die Stimmen nicht selbst ausgezählt, selbst wenn es prinzipiell denkbar wäre, dass ich das tun könnte. Ich weiß zwar erfahrungsbasiert, wen ich selbst gewählt habe. Ich weiß aber nicht autark, wie viele Stimmen die jeweiligen Parteien erhalten haben; das weiß ich hingegen vertrauensbasiert. Hier verläuft die Grenzlinie zwischen dem autarken und dem vertrauensbasierten Wissen.

3.2. Das Pervasivitätsmodell

In einem zweiten Modell gibt es keinen Bereich reinen Wissens und auch keine klare Demarkationslinie. Vertrauen und Wissen durchdringen einander vielmehr wechselseitig, sie haben eine verwickelte, verwachsene und intime Beziehung zueinander. Es gibt zwei Argumente, die für diese pervasive Beziehung von Wissen und Vertrauen sprechen.

Die Vertrauensgeladenheit von Wissen

Das erste Argument wurde von C.A.J. Coady (1992) entwickelt. Es zeigt auf, dass unsere eigene Erfahrung auf vielfache und nicht klar abgrenzbare Weise von dem, was wir durch andere ‚erfahren' haben, durchdrungen ist. Mit anderen Worten: Das Argument greift die essentielle Annahme der ersten Version an, dass es einen Bereich purer, *eigener* Erfahrung gibt, in dem andere *keine* Rolle spielen. Andere spielen grundsätzlich eine Rolle in dem, was wir erfahren.

Coady erläutert die Fraglichkeit dieser Annahme an dem Beispiel einer Person, die die britische Queen in Melbourne sieht. Gemäß dem Demarkationsmodell möchte man annehmen, es handelt sich dabei um eine *eigene* Erfahrung. Diese Person weiß qua eigener Kraft, dass die Queen in Melbourne ist; sie ist nicht von anderen, etwa Nachrichten abhängig, denn sie hat es mit eigenen Augen gesehen.

Coady argumentiert jedoch, dass die vermeintliche Unabhängigkeit der Erfahrung von anderen hier nicht besteht: die Beobachtung der Queen ist vielmehr durchsetzt von dem, was andere berichtet und vermittelt haben (ebd.: 147). So weiß die beobachtende Person etwa, wie die Queen aussieht, gar nicht qua eigener Erfahrung, sondern weil es ihr andere gesagt oder gezeigt haben – in der Form von Beschreibungen, von Fotos oder Filmaufnahmen, bei denen andere bezeugt haben: *das, was hier zu sehen ist, ist die Queen*. Ferner weiß die Person durch die Berichte anderer, wie ein solcher Besuch ungefähr aussehen wird; dass die Queen etwa von Leibwächtern, Angestellten und anderen Personen umgeben sein wird; würde die Person etwa eine andere Person sehen, die zwar wie die Queen aussieht, aber schlicht ihren Einkäufen im kleinen Eckladen nachgeht, würde sie kaum zur Überzeugung kommen: ‚Ich sehe die Queen!'

Für Coady zeigen diese Überlegungen, dass die Quellen unseres Wissens, nämlich Wahrnehmung, Erinnerung, Inferenzen und Bezeugungen nicht getrennt nebeneinander bestehen, sondern einander durchdringen.[7] Für unseren Zusammenhang kommt es auf folgenden Punkt an: die Erfahrung erweist sich als *vertrauensgeladen* (‚trust-loaden'), um hier den bekannten Terminus der Theoriegeladenheit von Beobachtung abzuwandeln.

Was Coady damit zeigt, hat unmittelbare Relevanz für die wissenschaftliche Praxis. Die wissenschaftliche Erfahrung im Rahmen von Beobachtungen, Messungen sowie Visualisierungen aller Art ist in einem hohen Maße dadurch geprägt, dass andere sie informieren und vorstrukturieren. Dies schließt nicht aus, dass Erfahrung die Rolle spielen kann, die ihr klassischerweise zugesprochen wird, etwa Hypothesen zu widerlegen. Doch haben dabei andere eine Stimme, denen wir vertrauen, wenn wir erfahren, was wir erfahren.

Das in Wissen eingekapselte Vertrauen

Das zweite Argument entwickeln wir ausgehend von einer Überlegung Bernd Lahnos (2002). Lahno geht es um das Verhältnis

von Evidenz und Vertrauen, wie es im Rahmen der epistemischen Ansätze entwickelt wurde. Gemäß dieser Ansätze gilt: Rationales Vertrauen *basiert* auf Evidenz. Umgekehrt bedeutet dies: Evidenz ist *unabhängig* von Vertrauen. Evidenz stellt das vertrauensfreie Fundament dar, von dem ausgehend eine Entscheidung darüber, ob anderen Vertrauen gewährt werden kann, getroffen werden sollte. Vertrauen ist eine Inferenz, die *abhängig* von dieser Evidenz ist; Vertrauen sollte entsprechend, so bereits Hume[8], nur in dem Maße gewährt werden, in dem Evidenz (für die Vertrauenswürdigkeit) gegeben ist. Erkennbar wird an dieser Konzeptualisierung, dass es wiederum einen Bereich des reinen Wissens geben soll.

Lahno bestreitet, dass es in vielen relevanten Fällen eine solche Unabhängigkeit gibt. Er erläutert dies am Beispiel eines Freundes, der eines Verbrechens beschuldigt wird. Die gegen diesen Freund vorgelegte Evidenz wird im Lichte des Vertrauens in den Freund betrachtet und bewertet. Lahnos Punkt ist folglich, dass diese Unabhängigkeit hier nicht besteht:

> „Sie werten die Informationen nicht einfach sachlich aus und bilden sich dann ein Gesamturteil. Die besondere Beziehung zu Ihrem Freund bringt Sie vielmehr dazu, die Informationen in einem ganz bestimmten Licht zu sehen, sie in Frage zu stellen." (Lahno 2002: 138)

Lahnos Gedanke ist, dass aus epistemischen Gründen Evidenz häufig nicht autark eingeschätzt werden kann.[9] Evidenz wird stattdessen im Licht bestehender Vertrauensverhältnisse bewertet. Zunächst könnte dieses Argument Lahnos so gelesen werden, dass eine Person, die einer anderen vertraut, eine psychische Voreingenommenheit hat, also einen Bias, der dazu führt, dass sie die Evidenz, die gegen die Person spricht, der sie vertraut, ausblendet oder nicht hinreichend wertet. Diese Lesart wäre verträglich mit dem Gedanken, dass die Evidenz gleichwohl autark von ihr bewertet werden *könnte*. Dies scheint uns aber nicht Lahnos Gedanke zu sein. Sein Punkt ist vielmehr, dass es in vielen Fällen *keine autarke Möglichkeit gibt*, Evidenz zu bewerten; man

kann sich diese Evidenz nicht eigenständig beschaffen, sondern ist von anderen abhängig, die diese (vermeintliche) Evidenz präsentieren. Ob man das von anderen Angebotene (Mitgeteilte) annimmt oder ablehnt, hängt davon ab, ob man ihnen vertraut oder misstraut.

Wir wollen nun einen Schritt weitergehen und einen neuen Punkt machen: Wenn wir Evidenz auf der Grundlage unseres Vertrauens in andere gewinnen, dann hilft uns diese Evidenz häufig weiter, um von ihr ausgehend neues Wissen zu erschließen. Betrachten wir dazu ein einfaches Beispiel: Eine Wissenschaftlerin erhält Messdaten von einem Kollegen. Sie verwendet diese im Rahmen eines computerbasierten Modells zur Initiierung einer Simulation. Auf diese Weise gewinnt sie eine Prognose über das mögliche Verhalten des Systems unter den gewählten Startbedingungen. Dabei vertraut sie dem Kollegen, der ihr die Messwerte zugesandt hat. Sie *vertraut* auf die Evidenz über den vorherigen Systemzustand. Dieses Vertrauen in das behauptete Wissen über vergangene Systemzustände ermöglicht es ihr, ihre eigene Forschung fortzuführen, in der sie sodann *eigenes Wissen* über künftige Systemzustände gewinnt. In diesem neu gewonnenen *eigenen* Wissen ist das *Vertrauen in ihren Kollegen* gleichsam eingekapselt. Die evidentielle Beweiskraft der Simulationsresultate wird u.a. dadurch gewonnen, dass der Datenquelle vertraut wird; *auf dieser Grundlage* wird dann eigene Evidenz generiert.

Anders als in Coadys Fall gilt: Die Beobachtung der Simulationswissenschaftlerin ist gewissermaßen autark. Die Wissenschaftlerin selbst beobachtet etwas an der Simulation. Der Punkt ist also nicht die Vertrauensgeladenheit ihrer Beobachtung (die sich darin manifestiert, dass das, was sie sieht, von anderen erlernt worden ist). Wir können sogar annehmen: Sie sieht etwas zum ersten Mal, das niemand vor ihr in einer solchen Simulation gesehen hat. Sie kann dies aber nur deshalb als aufschlussreiches Ergebnis einer Simulation sehen, weil sie den Daten, die sie zur Initiierung der Simulation verwandt hat, vertraut. Ihre Einsicht basiert auf (eingekapseltem) Vertrauen; Vertrauen in die Daten, die sie erst in die Lage versetzen zu beobachten, was sie beobachtet.

Wir haben damit zwei Argumente für die *Pervasivität* von Vertrauen und Wissen vorgestellt:

- die Vertrauensgeladenheit von Erfahrungen sowie
- das in Wissen eingekapselte Vertrauen.

Beide führen dazu, dass die Grenzen zwischen Vertrauen und Wissen nicht oder zumindest nicht eindeutig zu ziehen sind. Damit wollen wir nicht behaupten, dass Vertrauen und Wissen sich nicht begrifflich unterscheiden würden, sondern dass deren Beziehung verwickelter ist, als es gemäß der Demarkationsversion erscheint.

Diese intime Verwachsenheit von Vertrauen und Wissen hat Konsequenzen für die Versuche, anhand von Kriterien Wissensansprüche zu bewerten. Insbesondere in Fällen, in denen es darum geht, über die Expertise von Expertinnen zu entscheiden, zeigt sich, dass dies nicht bloß Vertrauen voraussetzt – in die Expertise von Experten –, sondern dass Vertrauen und Wissen einander durchdringen. Und zwar sowohl in Beziehungen zwischen Novizen und Expertinnen als auch zwischen Experten und anderen Expertinnen.

4. Der Ursprung des Erregers SARS-CoV-2: Die Labor-These

Wir kommen nun zu unserer Fallstudie, in der es um das Verhältnis von Vertrauen und Wissen geht. Unsere Frage ist die nach der Möglichkeit von Novizen, sich mittels Kriterien, wie denen von Goldman vorgeschlagenen, in einem Konflikt zwischen Expertinnen zu orientieren. In unserem Fall geht es um den Ursprung von SARS-CoV-2. Dieser Fall bietet eine hervorragende Basis für unsere Untersuchungsfrage:

Einander widersprechende Expertinnenurteile: In dieser Frage treffen u.a. zwei Positionen aufeinander: Expertinnen, die davon ausgehen, dass der Covid-19-Erreger natürlichen Ursprungs ist; sowie Expertinnen, die annehmen, dass sein Ursprung in einem

Labor (in Wuhan) liegt. In diesem Fall gibt es also die Positionen E_1 und E_2, die einander widersprechen. Die Positionen unter den Expertinnen sind vielfältiger, als es diese beiden groben Charakterisierungen nahelegen. Zum einen kommt die Beurteilung der vorliegenden Evidenz für jede der beiden Thesen in allen möglichen Schattierungen vor: von Experten, die sie für (relativ) zwingend halten, bis zu solchen, die sie nicht bloß als vorläufig, sondern insbesondere noch als unzureichend für irgendwelche Schlüsse erachten und sich daher eines Urteils enthalten. Gleichwohl nimmt der Widerstreit der beiden genannten polaren Positionen in der Öffentlichkeit einen zentralen Raum ein.

Engagiert beteiligte Öffentlichkeit: Neben den einander z.T. heftig widersprechenden Expertinnen gibt es eine engagierte Öffentlichkeit – Politikerinnen, Journalisten, Bürgerinnen usw. – die als Novizen mit dieser Debatte konfrontiert werden und versuchen, sich in ihr zu orientieren. In Meinungsbeiträgen, Berichten, Foren und Blogs, auf Podien und in Pressekonferenzen ist über die Frage nach dem Ursprung des Virus gestritten worden.

Politik: Diese Debatte ist politisch extrem stark aufgeladen (worden). Dazu trägt bei, dass die Frage nach dem Ursprung des Virus untrennbar mit der Frage nach der Verantwortung für die Pandemie verbunden ist. Die Frage nach dem Ursprung betrifft die Frage, wer (mit)verantwortlich für Millionen gestorbener Menschen sowie für all die weiteren Folgen der Pandemie und ihrer Bekämpfung ist. Die Frage nach dem Ursprung hat zudem eine geopolitische Bedeutung: Im Hintergrund steht der Konflikt zwischen den USA und China sowie der Kampf darum, das jeweilige politischen System als dem anderen überlegen darzustellen.

Wissenschaft & Politik: Ferner ist die Frage von Relevanz für das Vertrauen in die Wissenschaft. Würde sich ergeben, dass das Virus seinen Ursprung im Labor hätte, hätte dies ganz andere, vermutliche gegenteilige Folgen für die dort betriebene Art der Forschung, als wenn es natürlichen Ursprungs wäre. In den USA geht es darüber hinaus um die Rolle, die Wissenschaft und insbesondere wissenschaftliche Expertise in der Politik und in Fragen gesellschaftlicher Orientierung spielen sollen. Der dortige Streit

um die Rolle von Wissenschaft und wissenschaftliche Expertise manifestierte sich in den zu Coronazeiten geäußerten und auf T-Shirts gedruckten Parolen wie „Trust the Science!" oder „In Fauci We Trust".

Politische Identität: In den USA, in Ansätzen aber auch in Deutschland, sind die Positionen in der Debatte um den Ursprung des Covid-19-Erregers zugleich mit politischen Positionen in der Bewertung der Maßnahmen zur Eindämmung von SARS-CoV-2 identifiziert worden, was nicht zuletzt daran lag, dass sich Donald Trump relativ frühzeitig dezidiert für die Laborursprungsthese ausgesprochen hatte. Eine Äußerung für oder wider die Lableak-These wurde damit zugleich als Enthüllung der politischen Identität der Personen verstanden.

Desinformation: Schließlich ist dieser Fall auch deshalb besonders für unsere Untersuchung geeignet, da sich um ihn konspirative Theorien (über Wissenschaftler, die im Geheimen agieren und die Öffentlichkeit zu täuschen versuchen; über Institutionen und Medien wie auch über Regierungen, die federführend oder assistierend daran mitwirken) ranken und er selbst zahlreiche Desinformationsversuche evozierte.

Die Debatte weist, wie zu sehen ist, mehrere Schichten auf und ist mit großer Intensität geführt worden. Zu ihr wäre erheblich mehr zu sagen, als wir an dieser Stelle ausführen können. Ein Hinweis vorab ist jedoch unabdingbar: die These, nach der der SARS-CoV-2 aus einem Labor (in Wuhan) entstammt, ist vieldeutiger, als es zuweilen erscheint. Sie kann u.a. meinen:

1. Das Virus ist aus einem Labor nach draußen gelangt;
   (a) durch einen Unfall
   (b) beabsichtigt
2. Das Virus wurde im Labor gezielt verändert;
   (a) etwa um es zu erforschen und vorsorglich gegen eine mögliche künftige Pandemie gewappnet zu sein;
   (b) um eine Biowaffe herzustellen.

Die These von einem Laborleak meint notwendigerweise, dass Aussage 1 erfüllt ist, zuweilen wird darüber hinaus auch Aussage

2 eingeschlossen (in Variante a oder b). Um eine gewisse Übersicht über die Debatte zu bieten, stellen wir einige Passagen ihrer zeitlichen Entwicklung vor. In keiner Weise ist die folgende Übersicht als nur annähernd vollständig zu betrachten: Nicht nur fehlen viele Ereignisse; zu jedem genannten Ereignis gibt es viele weitere Informationen, Einschätzungen und vor allem Kontexte, die fehlen.

Zur Vorgeschichte von SARS-CoV-2 gehört natürlich SARS-CoV-1. Erste Fälle einer Infektion mit SARS-CoV-1 wurden Ende 2002 in Südchina bekannt. Die Sterblichkeitsrate dieses *S*evere *A*cute *R*espiratory *S*yndrome (SARS) war deutlich höher als bei vielen anderen Atemwegserkrankungen. 2005 wurden Coronaviren in Fledermäusen nachgewiesen. Vermutet wird, dass das Virus von Fledermäusen über einen Zwischenwirt auf Menschen übertragen wurde. Zwar konnte es relativ schnell eingedämmt werden (auch weil es weitaus weniger infektiös als der Covid-19-Erreger war), jedoch wurde es als eine mögliche Gefahr für eine künftige Pandemie erkannt. In diesem Zusammenhang wurde die Suche nach Virusvarianten in Fledermäusen intensiviert. 2013 wurde die Coronavirussequenz RaTG13 vom Wuhan Institute of Virology (WIV) entdeckt. Es handelt sich um das Virus, das die damals bekannteste genetische Ähnlichkeit (ca. 96 %) mit dem Coronavirus, das Covid-19 verursacht, aufweist.

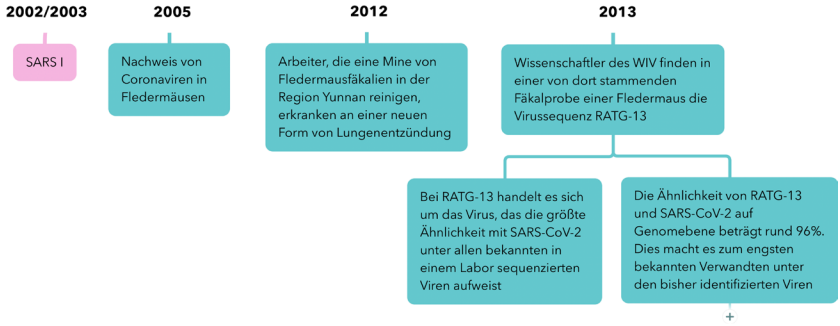


Abbildung 7.2: Vorgeschichte von SARS-CoV-2.

Weniger als drei Monate nachdem die ersten Fälle der von SARS-CoV-2 verursachten neuen Krankheit *Covid-19* bekannt wurden, erschien bereits ein einflussreicher Artikel mit dem Titel „The proximal origin of SARS-CoV-2" (Andersen u.a. 2020) in der Zeitschrift *Nature Medicine*, in dem mit Blick auf die damalige Kenntnislage die Lableakthese als unplausibel bewertet wurde; insbesondere der Gedanke einer gezielten gentechnischen Gestaltung, aber im Kern auch die schwächere These eines Austritts aus dem Labor werden dort als nicht durch die vorliegende Evidenz gestützt betrachtet.[10] Als im April 2020 der damalige US-Präsident Donald Trump und sein wissenschaftlicher Berater Anthony Fauci, Direktor des *National Institute of Allergy and Infectious Diseases* (NIAID) und Mitglied der Coronavirus-Taskforce der US-Regierung, vor die Presse traten, wurde an beide die Frage nach dem Ursprung von SARS-CoV-2 gestellt. Trump überließ Fauci die Antwort. Dieser verwies auf den „Proximal Origin"-Artikel mit den Worten, dass dort der natürliche Ursprung in der Übertragung vom Tier auf den Menschen als plausibel betrachtet wurde. Im Anschluss bot er der Presse an, den entsprechenden Artikel zur Verfügung zu stellen.[11] Als Trump wenige Tage später allein vor die Presse trat und erneut nach dem Ursprung des Virus sowie danach gefragt wurde, ob er Evidenz gesehen habe, die für einen so genannten Lableak spricht, antwortete er nun: „Yes, I have".[12] Er könne die von ihm eingesehene Evidenz aber nicht mit der Öffentlichkeit teilen; daraufhin beschimpfte er die World Health Organization (WHO) und warf ihr vor, eine PR-Abteilung Chinas zu sein.

Ein Team der WHO reiste im Januar 2021 u.a. nach Wuhan, um den Ursprung des Virus zu untersuchen. Im Untersuchungsbericht wurden vier mögliche Ausbreitungswege („pathways") des Virus diskutiert und bewertet. Als wahrscheinlich bis sehr wahrscheinlich wurde eine Übertragung des Virus von Fledermäusen über einen Zwischenwirt auf Menschen betrachtet. Dies würde etwa zur These passen, dass der Huanan Seafood Market zentral für den Ursprung der Pandemie war. Als ein „extremely

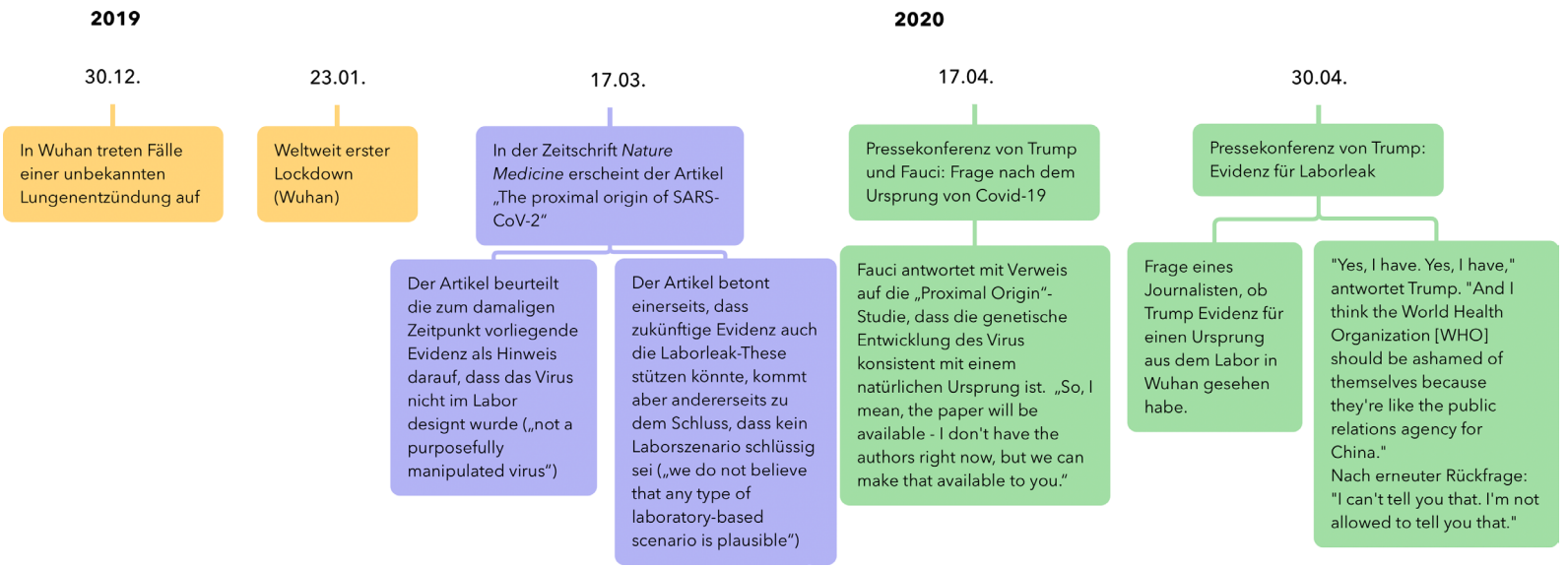


Abbildung 7.3: Die wissenschaftliche und politische Kontroverse beginnt.

unlikely pathway" wurde dagegen der Laborunfall bewertet.[13] Am 07. März 2020 erschien ein Aufruf von Wissenschaftlern in der Zeitschrift *The Lancet*, in dem u.a. der solidarische Zusammenhalt mit chinesischen Wissenschaftlern gefordert wurde. Darüber hinaus hieß es: „The rapid, open, and transparent sharing of data on this outbreak is now being threatened by rumours and misinformation around its origins. We stand together to strongly condemn conspiracy theories suggesting that COVID-19 does not have a natural origin." (Calisher u.a. 2020: e42)[14] Ein im Mai erschienener Brief von Wissenschaftlerinnen in der Zeitschrift *Science* rief dagegen bereits im Titel zu (mehr) Offenheit auf: „Investigate the origins of COVID-19". Kritisiert wurde insbesondere der Untersuchungsbericht der WHO. Im Brief heißt es: „The information, data, and samples for the study's first phase were collected and summarized by the Chinese half of the team; the rest of the team built on this analysis." (Bloom u.a. 2021: 694)[15] Die beiden zentralen Ursprungsthesen (Labor vs. Natur) wären in dem Bericht zudem nicht ausgewogen behandelt worden.[16] Man sieht an diesen beiden Stellungnahmen, wie sich innerhalb der wissenschaftlichen Community eine nicht nur wissenschaftliche Debatte entwickelt; es geht darin nicht ‚bloß' um die Interpretation von Daten und die Bewertung von Evidenz für unterschiedliche Thesen, sondern auch um Desinformation, (politischen) Zugang zu Daten, politischen Bias usw.

Neu angefacht wurde die wissenschaftliche und insbesondere politische Debatte, als im September 2021 ein bei der DARPA[17] eingereichter Forschungsantrag geleaked wurde (vgl. Hibbett und Lerner 2021).[18] Antragsteller war die NGO EcoHealth Alliance, deren Präsident Peter Daszak als federführender Projektleiter (PI) vorgesehen war. In dem eingereichten Antrag „DEFUSE: Defusing the Threat of Bat-borne Coronaviruses" wurde ein Plan entworfen, die Gefahren von Coronaviren, die aus Fledermäusen stammen, zu untersuchen, um Gegenmaßnahmen zu entwickeln. Dazu gehörte auch der Vorschlag, im Labor Virusvarianten künstlich zu erzeugen, indem genetische Abschnitte aus verschiedenen Coronaviren miteinander kombiniert werden sollten (*chimäre Coronaviren*). Dabei sollten einzelne Gene

oder Gensegmente aus einem Coronavirus durch das Segment eines anderen ersetzt werden. Erforscht werden sollte dadurch, ob bestimmte Virenvarianten eine erhöhte Infektiosität für Menschen aufweisen könnten. Im Antrag wurde auch vorgeschlagen, eine humanspezifische Furin-Spaltstelle in die Spike-Proteine bestimmter Coronaviren einzufügen. Furin-Spaltstellen können es dem Virus ermöglichen, menschliche Zellen effizienter zu infizieren, was es potenziell leichter übertragbar macht (ein Ergebnis, das einem „Gain-of-Function" entspricht). Tatsächlich ist SARS-CoV-2 insofern ungewöhnlich, als es eine Furin-Spaltstelle besitzt, die andere bekannte Coronaviren der Gruppe, zu der SARS-CoV-2 gehört, nicht aufweisen (es gibt aber andere bekannte Coronaviren, die Furin-Spaltstellen besitzen).

Die Forschung sollte in Kooperation mit dem Wuhan Institute of Virology (WIV) in einem Labor in North Carolina durchgeführt werden. Der 2018 eingereichte Antrag war letztlich von der DARPA abgelehnt worden. Es gab dennoch vier Gründe, weshalb der Leak besondere Aufmerksamkeit erregte: (1) Die Furin-Spaltstelle, (2) dass der Antrag erst zu diesem Zeitpunkt (durch einen Leak) bekannt wurde; (3) dass er von Peter Daszak als PI verfasst wurde und (4) die Forschung in Kooperation mit dem in Wuhan ansässigen WIV erfolgen sollte. Daszak war Mitglied der Gruppe, die im Auftrag der WHO den Ursprung von SARS-CoV-2 untersucht hatte (er war Mitglied des China-Teams; der Untersuchungsbericht war, wie oben dargestellt, von Wissenschaftlerinnen in dem Brief, der in der Zeitschrift *Science* erschien, kritisiert worden). Unter anderem dieser Interessenskonflikt rief Misstrauen hervor. Zudem wurden Emails öffentlich zugänglich, die zeigten, dass Daszak seine Rolle als Initiator und zentraler Autor der Stellungnahme, die im März im *Lancet* erschienen war, geheim zu halten überlegte.[19]

Wir selbst wollen und können uns nicht zu dieser Frage verhalten oder die Evidenz bewerten, da uns die Expertise hierzu fehlt. Das Anliegen dieses Abschnitts ist es, eine knappe Übersicht über die öffentliche Debatte zu geben. Das ist auch deshalb so schwer, da jede weitere Information, die wir geben könnten, das Bild (gerade

für Laien) verändern mag.[20] Und jede (erforderliche) Kommentierung ebenso. Gleichwohl sollte auch nicht der Eindruck entstehen, dass es ein Gleichgewicht unter den Expertinnen mit Blick auf die Kernfrage *Natur oder Labor* gibt. Es wurden mehrere weitere Studien vorgelegt. Bis heute scheint es eine Mehrheit von Wissenschaftlern zu geben, die die bislang vorliegende Evidenz so lesen, dass sie einen natürlichen Ursprung des Covid-19-Erregers nahelegt.[21] Aktuelle Zeitungsberichte deuten aber an, dass es weitere Evidenz gibt, die für die Laborleak-Hypothese spricht und auch von renommierten Wissenschaftlern (vorsichtig) als solche gewertet wird.[22]

Gegen die Idee, dass das Virus gentechnisch im Labor gestaltet wurde, spricht unter anderem, dass der Virenstamm, der die größte Ähnlichkeit mit SARS-CoV-2 aufweist, nämlich die vom WIV in Fledermäusen entdeckte Sequenz RaTG13 eine genetische Ähnlichkeit von 96% aufweist. Für Laien mag dies als eine beeindruckend hohe Ähnlichkeit erscheinen, für Expertinnen bedeuten 96% aber, dass der Abstand zwischen beiden Sequenzen immens ist: zu groß insbesondere, als dass er als Ausgangssequenz für eine gentechnische Modifikation in Frage kommt. Ferner wäre es auch fraglich, warum die immense Arbeit an weiteren Veränderungen durch gezielte technische Modifikationen unternommen worden wäre, die für die Gestaltung eines Coronavirus mit erhöhter Infektiosität für Menschen irrelevant sind. Schließlich spricht gegen ein Design der humanspezifischen Furin-Spaltstelle, dass diese keineswegs optimal erfolgte. Hätte man diese designen wollen, wären andere Optionen besser gewesen. Gegen die These, dass das Virus aus dem Labor entwichen ist (ohne dort gentechnisch verändert worden zu sein), sprechen zudem die Orte, an denen die bekannten frühen Fälle infizierter Personen gefunden wurden; ferner die verschiedenen Linien des Virus sowie die viralen Spuren in Tieren, die man vom Markt gewonnen hat; gleichwohl ist gerade mit Blick auf letzteren Punkt eingewandt worden, dass es hier einen Bias gegeben haben mag, da andere Orte nicht oder nicht mit der gleichen Sorgfalt in Betracht gezogen worden seien.[23]

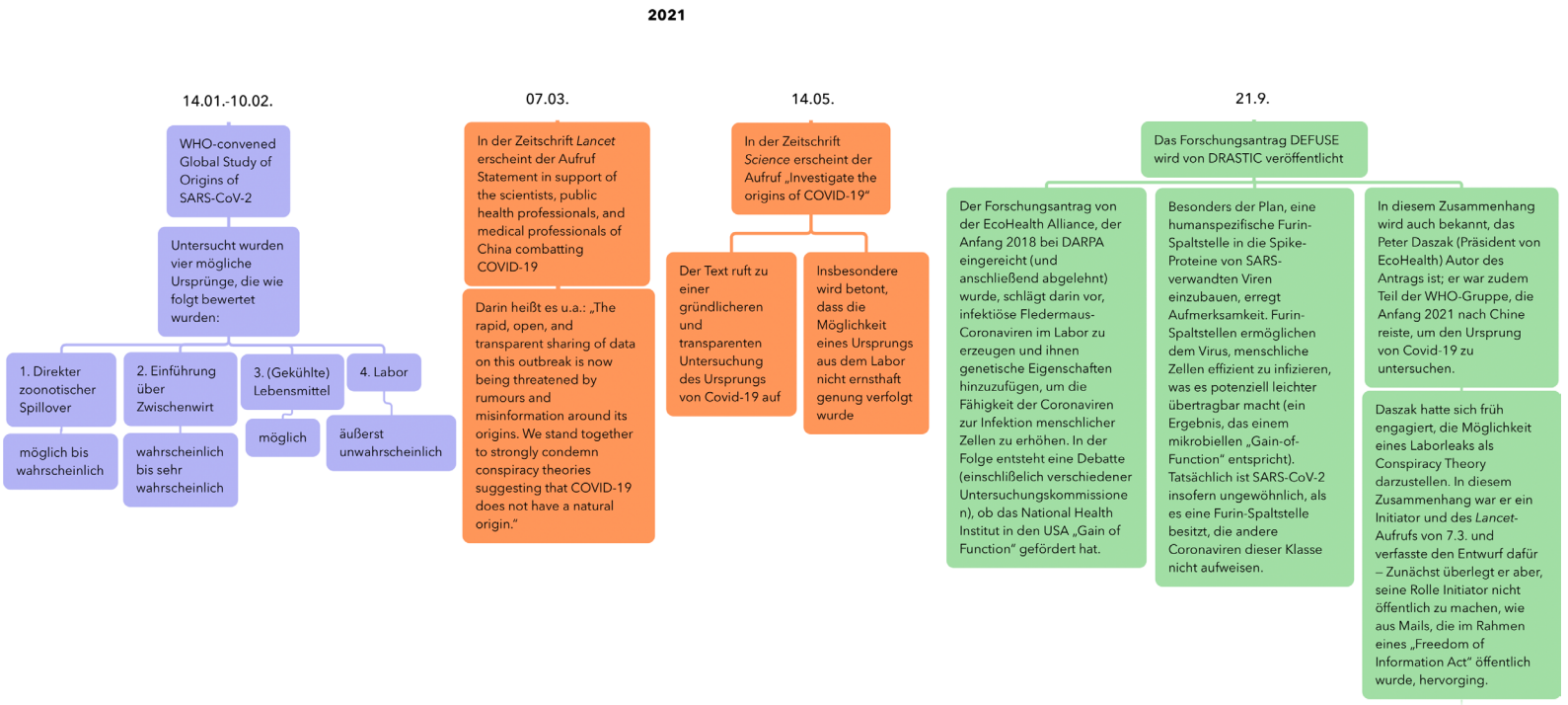


Abbildung 7.4: Intensivierung der Kontroverse.

5. Wie Novizen sich in der Frage über den Ursprung von SARS-CoV-2 orientieren

Bislang haben wir primär die sich entwickelnde Kontroverse in der Diskussion unter Expertinnen vorgestellt. Wir verlassen nun diese Ebene und fragen stattdessen: Wie können sich Novizen in dieser Kontroverse orientieren? Wie können sie sich vernünftig zu ihr verhalten und entscheiden, welchen Expertenaussagen sie vertrauen können?

Einen guten Ausgangspunkt, um dieser Frage nachzugehen, stellen Kommentare dar, die Leser zu einem als Meinungsbeitrag rubrizierten Text von Alina Chan in der *New York Times* verfasst haben. Chan ist eine Molekularbiologin und arbeitet als Postdoktorandin am mit dem MIT und der Harvard Universität affiliierten Broad Institute, wo sie unter anderem das Projekt *COVID-19 Coronavirus Genetics* leitete, ein Open-Source-Projekt, das die genetische Entwicklung von SARS-CoV-2 nachzeichnet. Sie ist zugleich eine der prominentesten Fürsprecherinnen der Laborleakthese. Bereits seit 2021 hat sie hierzu öffentlich Position bezogen und verteidigt die These auch heute noch in verschiedenen Medienformaten. In ihrem Text „Why the Pandemic Probably Started in a Lab" vom 03. Juni 2024 fasst Chan fünf Hauptargumente zusammen, die aus ihrer Sicht für die Laborhypothese sprechen – und damit gegen die Markthypothese (vgl. Chan 2024). Diese Argumente sind recht unterschiedlicher Natur. Sie beziehen sich: auf die räumliche Nähe zwischen dem Ausbruchsort der Pandemie und dem Standort eines der führenden Labore für SARS-verwandte Viren (1. Argument); den erwähnten DEFUSE-Forschungsantrag (2. Argument); die im Vergleich zu US-amerikanischen Standards niedrige Biosicherheitsstufe, unter der im WIV geforscht wurde (3. Argument); Daten zur Evolution des Virus, die nahelegen, dass die ersten am Markt infizierten Personen vermutlich nicht die ersten Fälle von Covid-19 bei Menschen waren (4. Argument); und die fehlende Schlüsselevidenz, die Infektionen bei Menschen mit den Infektionen bei Tieren verbinden würde (wie etwa noch nicht gefundene direkte Vorgängervarianten

des Virus in Tieren, wie man sie bei SARS-CoV-1 im Jahr 2022 und bei MERS im Jahr 2012 gefunden hatte) (5. Argument).

Der Beitrag verzeichnete nach Schließung der Kommentarspalte erstaunliche 1.597 Kommentare. Vereinzelt kommt es dazu, dass Kommentatoren auf andere reagieren und sich kleine, fokussiertere Diskussionen entspinnen. Die Beiträge sind überraschend wenig repetitiv und bringen eine Vielfalt an Aspekten in die Diskussion ein. In ihnen werden unterschiedliche Kriterien angewandt, mit denen die Kommentierenden für oder wider die Glaubwürdigkeit der Autorin argumentieren – aber, wie wir unten sehen werden, auch direkte Argumente erster Stufe anbringen, um sich zu der Frage nach dem Ursprung des Virus zu verhalten. Wohlgemerkt wollen wir damit nicht behaupten, dass die Kommentare hier ein *repräsentatives* Meinungsbild abgeben. Ihre Funktion für unsere folgende Untersuchung liegt vielmehr darin, die allgemeinen Orientierungspunkte von Novizen anhand von herausgegriffenen realen Beispielen zu illustrieren.

Damit wollen wir den Ansatz von Goldman erweitern. Goldman selber verstand seine fünf Kriterien als tentative Schritte auf einem noch von der Philosophie weitgehend unbehandelten Problemfeld. Auch betonte er, dass seine Kriterienliste keinen Anspruch auf Vollständigkeit hat. Sein Ansatz wurde inzwischen erweitert und verfeinert, etwa in Beiträgen von Oliver Scholz (2009), Elizabeth Anderson (2011) und Gloria Origgi (2022). Wir werden an diese Erweiterungsversuche anknüpfen und noch weitere Dimensionen zur Diskussion stellen, anhand derer sich Novizen in Bezug auf eine wissenschaftliche Aussage epistemisch orientieren können.

5.1 Motive oder Charaktereigenschaften der vermeintlichen Expertin

Goldman selber brachte mit seinem vierten Kriterium der (finanziellen) Interessen bereits die Ebene der Motive ins Spiel. Dieser Ansatz scheint aus einem grundlegenden Gedanken heraus

plausibel: Wenn es um die Frage geht, ob einer Expertin vertraut werden kann, ist nicht nur die Bewertung ihrer Expertise, gedacht als Kombination von Fachwissen und kognitiven Fähigkeiten, entscheidend. Essenziell ist auch, dass die Expertin darauf bedacht ist, die Wahrheit zu erkennen und sie mitzuteilen. Neben der Fertigkeit, die Validität wissenschaftlicher Aussagen einschätzen zu *können*, sollte sie also auch dazu disponiert sein, die epistemische Validität ernsthaft und gewissenhaft einschätzen und kommunizieren zu *wolle*n.

Auf dieser Ebene können die von Goldman genannten Anreizstrukturen, die Motivlage von Expertinnen ungünstig beeinflussen. Darunter fallen etwa die genannten finanziellen Interessen, denen laut Goldman die von der Industrie vorgenommene Forschung unterliegen kann.[24] Man kann aber auch auf Charakterdispositionen einzelner Expertinnen rekurrieren, insbesondere auf deren epistemische Tugendhaftigkeit.[25] Epistemische Tugenden sind charakterliche Dispositionen von Personen, ihre Erkenntnisprozesse verantwortungsvoll vorzunehmen. Dazu zählen etwa intellektuelle Aufgeschlossenheit und eine Disposition zum kritischen Hinterfragen sowie zu gewissenhaften Untersuchungen. In ihnen kommt ein Bemühen zum Ausdruck, ernsthaft an der Wahrheit interessiert zu sein. Epistemische Tugenden stellen zwar keine Fertigkeiten oder Wissen dar[26], sie sind also kein Teil der Expertise im goldmanschen Sinne. Gleichwohl erhöhen sie die Vertrauenswürdigkeit einer Expertin.

Zu den entscheidenden Tugenden eines Experten gehören aber auch kommunikative Tugenden wie die Ehrlichkeit der Person (Anderson 2011: 146). Ein Experte könnte zwar zu den richtigen Überzeugungen kommen, in der Öffentlichkeit jedoch unaufrichtig sein, lügen oder sich irreführend oder in verzerrender Weise äußern. Neben der Ehrlichkeit einer Person scheint zudem auch eine gewisse *kommunikative Gewissenhaftigkeit* eine entscheidende Charaktereigenschaft glaubwürdiger Experten darzustellen: Damit soll die Disposition gemeint sein, sich erst dann zu einem wissenschaftlichen Sachverhalt zu äußern, wenn man sich hinreichend bemüht hat, die Sache gut einschätzen zu können.

Ein Anzeichen, dass eine Expertin sich eher unbedacht und voreilig in der Öffentlichkeit äußert, wäre daher ein Grund, ihr nicht zu vertrauen.

Die entscheidende Anschlussfrage ist nun aber: Woran können Novizen motivationale oder charakterlichen Eigenschaften von Expertinnen erkennen? Neben Goldmans Idee, ökonomische Interessen in den Blick zu nehmen, lassen sich noch eine ganze Reihe weiterer Gesichtspunkte ins Feld führen. Die Kommentare zum *New York Times*-Beitrag verweisen für den Fall der Laborhypothese auf ganz andere Kriterien, als sie Goldman diskutiert (finanzielle Interessen scheinen interessanterweise nur eine untergeordnete Rolle zu spielen). Hierzu nur zwei Punkte:

Zunächst ist bemerkenswert, dass des Öfteren der Anspruch erhoben wird, epistemische Tugenden oder Laster an der Erscheinungsweise der entsprechenden Personen direkt ablesen zu können. Exemplarisch dafür ist etwa der folgende Kommentar: „As a layman, what I find most convincing is the humility of those who have steadfastly insisted that, absent any hard evidence to the contrary, COVID *may* have come from a lab leak. That open-mindedness stands in stark contrast to the shrill certitude of those who denounce any questioning of the origin to be a 'right-wing conspiracy theory'" (Bill Bright (NY, NY), June 3). Anscheinend wird hier der Anspruch erhoben, epistemische Tugenden wie Bescheidenheit (*humility*) sowie Vorsicht und Offenheit in der Urteilsbildung *auch* – wie im Kommentar hervorgehoben wird – *als Laie* direkt am kommunikativen Verhalten der entsprechenden Akteure ablesen zu können. Ebenso ließen sich Laster, wie „schrille Gewissheit", die nicht einmal mehr gewisse Fragen zulasse, unmittelbar erkennen.

Zweitens ist bemerkenswert, wie häufig in Bezug auf unser Fallbeispiel *politische* Motive aufgerufen werden. Expertinnen werden demnach, politische (und nicht unbedingt finanzielle) Motive unterstellt – was wiederum als ein Faktor präsentiert wird, der an ihrer Glaubwürdigkeit kratzt und dazu führt, dass ihnen nicht vertraut wird. Ein Kommentator deutet etwa an, dass hinter Chans Meinungsbeitrag, tatsächlich eine politische Agenda stecke.

So hält er mit Verweis auf eine im Zeitraum der Veröffentlichung stattfindende Anhörung Anthony Faucis in einem Unterkommittee des Repräsentantenhauses fest: „Dr. Chan's view is the minority one. Odd that it is appearing at the same time Dr. Fauci is testifying. The republican obsession with proving the virus started in a lab has overtaken the more important matter of what future preventative matters should we be taking as a nation." (Michael J. McManus (South Bend, Indiana), June 3). Andere Kommentare sind noch deutlicher: „ ...this piece, coming in the context of yet another GOP kangoroo court today with Anthony Fauci suggests that Chan is playing politics not science and has exhausted her credibility with me" (JETIII (Portland), June 3). Ein weiterer Kommentar verweist auf den politischen Hintergrund des Ko-Autors des Buches von Alina Chan, Matt Ridley, der auch Politiker der britischen Conservative Party ist: „Ridley appears to be a libertarian skeptic of climate change who has been promoted by Rupert Murdoch. He presided over a failed bank that was then bailed out by the UK government. It seems plausible that Matt Ridley and Alina Chan have a right-wing political agenda" (Thomas (Boston), June 3). Die Beispiele zeigen die vielfachen Versuche, mal mehr, mal weniger plausibel, Indizien für (vermeintliche) politische Motive aufzufinden, um damit Vertrauens- bzw. Misstrauenseinstellungen zu begründen.

5.2 Das Medium der Veröffentlichung

Unseres Erachtens kann ein entscheidender blinder Fleck in der Diskussion um das Novize-2-Experten-Problem in der Rolle des *Mediums*, in dem eine Expertenaussage vorkommt, gesehen werden.[27] Tatsächlich scheint dies ein besonders relevanter Orientierungspunkt für Expertenvertrauen zu sein. Ob eine Person als Experte in der *Tagesschau* auftritt und dort etwa die weitestgehende Sicherheit von Impfstoffen garantiert oder sich auf seinem persönlichen Youtube-Kanal als selbsternannte Experten dazu äußert, ob eine Person von Joe Rogan in seinem Podcast als

Experte in dieser Frage vorgestellt wird oder sich auf dem rechtsgerichteten Medienportal *InfoWars* auftritt – der mediale Kontext der Äußerung scheint eine immense Bedeutung für unsere Einschätzung der Glaubwürdigkeit der Person zu haben.

Bei Goldman drängt sich dieser blinde Fleck frappant auf: So setzt er für das Szenario des Novize-2-Experten-Problems schlichtweg voraus, dass einer Novizin zwei *vermeintliche (putative)* Expertinnen gegenüberstehen. Offen bleibt, wie es dazu kommt, dass sie sich begegnen, und wie die Novizin dabei zu der Erwägung gelangt, dass es sich überhaupt um Expertinnen handeln *könnte* – deren Expertenstatus in der Folge zu ermitteln überhaupt erwägenswert und nicht schlichtweg abwegig wäre. Welcher andere Fall wäre hier vorstellbar, als dass der Novizin die beiden Personen in einem medialen Kontext begegnen, in dem ihr beide eben als vermeintliche Expertinnen *präsentiert* werden? – Sei es, da sie als Expertinnen in einem Interview befragt werden, auf einer Pressekonferenz sprechen oder zu einem Streitgespräch miteinander eingeladen sind.

Die Erklärung, warum Medien einen Orientierungswert in der Frage haben, ob wir geneigt sind, Expertinnen zu vertrauen, liegt darin, dass wir hinter den jeweiligen Medienformaten bestimmte Personen und institutionelle Gefüge annehmen, die redaktionelle Entscheidungen treffen. Darunter fällt auch die Auswahl der Personen, denen man eine öffentliche Bühne gibt. Im besten Fall erfüllen diese institutionellen Arrangements die Funktion von Gatekeepern, sorgen also für die epistemische Qualität der dort verbreiteten Inhalte (Mößner 2018: 214f.).

Auch in den *New York Times*-Kommentaren kommen mehrfach Bezüge zum Publikationsorgan ins Spiel, um die Glaubwürdigkeit der Autorin und anderer Experten zu verhandeln. Häufig wird als ein Grund für Skepsis angeführt, dass die Autorin ihre Argumente nicht in einer wissenschaftlichen Fachzeitschrift publiziert. Exemplarisch dafür fragt ein Kommentator: „Why did the author skip normal scientific process and decide to publish the article in a newspaper?" (JonShow (California), June 3) Die Praxis des Peer-Review wird in diesem Zusammenhang oft als ein ebensolcher

institutionalisierter Mechanismus ins Spiel gebracht, der der Qualitätssicherung des Veröffentlichten dient[28] – dass er hier ausbleibt, scheint ein entscheidendes Argument zu sein, um zumindest misstrauisch bezüglich der Aussagen von Chan zu sein.[29]

5.3 Die Herkunft der Daten

Novizen haben zwar *per definitionem* nicht die Expertise, um wissenschaftliche Argumente nachzuvollziehen. Aber sie haben die Möglichkeit, sich über den Ursprung der Daten zu informieren, auf die sich Expertinnen in ihren Analysen beziehen – und mögen bereits an dieser Stelle Skepsis anmelden. Dies bringt einen Aspekt mit ins Bild, der in der üblichen Betrachtung der Frage, ob Novizen Experten vertrauen können, schnell droht aus dem Blick zu geraten: Auch Expertinnen sind keine rein autarken Erkenntnissubjekte, sondern wesentlich darauf angewiesen, ihren Kolleginnen zu vertrauen (vgl. Kaminski 2024).

In Bezug auf unser Fallbeispiel kann man hierfür ein eindrückliches Beispiel nennen: Eine Schlüsselstudie in der Debatte war das oben genannte „Proximal Origin"-Paper, das von einer Gruppe Virologen im März 2020 in *Nature Medicine* veröffentlicht wurde (vgl. Anderson u.a. 2020). In der Studie argumentieren die Autoren nicht nur, dass SARS-CoV-2 Eigenschaften hat (namentlich die RBD- und die Furin-Spaltstelle), die nach den bekannten Verfahren so nicht im Labor hergestellt würden. Zugleich bezieht sich die Studie aber auch auf die Sequenz RaTG13. Diese ist die Genomsequenz einer in Fledermäusen gefundenen Virusprobe aus dem WIV in Wuhan, die SARS-CoV-2 am nächsten kommt. Die Autoren argumentieren diesbezüglich, dass es unwahrscheinlich sei, dass SARS-CoV-2 aus dieser Virusvariante im Labor hätte erzeugt werden können. Wahrscheinlicher sei im Vergleich vielmehr ein natürlicher Ursprung.

In einem Podcast der BBC zur Frage nach dem Virusursprung unterhält sich der Journalist und ehemalige Chinakorrespondent[30] John Sudworth mit einem der Autoren der Studie Robert

F. Garry. Sudworth stellt Garry die Frage, wie er sich eigentlich sicher sein kann, dass die von Shi Zhengli bereitgestellte Probe RaTG13 tatsächlich diejenige war, die SARS-CoV-2 am nächsten kommt, und Zhengli nicht vielmehr weitere Proben vorenthalten habe. Sudworth kommentiert diesen Gedanken zunächst, wie folgt: „And didn't it all once again, I found myself wondering, come down to the question of trusting information coming out of China? What if the WIV did have a closer virus to SARS-CoV-2, something other than RaTG13? Would a prestigious, national-level, Communist Party-controlled laboratory be likely or able to tell us?" (Sudworth 2023) Diese Frage legt das Vertrauen von Experten untereinander frei. In der Folge ergibt sich nun ein Dialog zwischen Sudworth und Garry:

> Garry: „You would have to assume, and a lot of people are trying to make this assumption that they had another virus, that they just weren't disclosing or telling a path. For me, that gets into kind of the area of conspiracy theory, right?"
> Sudworth: „But doesn't that mean that proximal origins' depends to some extent on simply trusting the WIV or trusting Xi Jinping?"
> Garry: „Well, I think we have to believe the scientific literature. I mean, I think there is some trust. I mean, I've heard Shi Zhengli give talks and things. She's clearly a brilliant scientist. If you want to believe that, you know, she's sort of the equivalent of an evil Dr. No-type of person that could hide this and, you know, do all this surreptitious work, I can't help you. All I can tell you is, yeah, I kind of trust her. She said she didn't have it. I believe her." (Episode 2: A Question of Trust, 25:05)

Garry hätte vielleicht, wie es andere gemacht haben, darauf verweisen können, dass es vor Beginn der Pandemie keinen Grund gegeben hätte, diese Sequenzen zu verbergen. Er hätte also Argumente vorbringen können, die nicht bloß auf Vertrauen beruhen. Gleichwohl dürfte es schwer von der Hand zu weisen sein, dass Vertrauen unter den Wissenschaftlerinnen eine Rolle spielt.[31]

Dieser Fall steht exemplarisch für eine generellere Schwierigkeit: Inwiefern kann Daten, die aus China kommen, vertraut werden?

Im Lichte dieser Frage ist es nicht verwunderlich, dass in den Kommentaren zum *New York Times*-Artikel Misstrauen in Experten oft mit „Misstrauen gegenüber China" begründet wird: „the WHO report that looked like a cover-up (superficially considered the lab but only via the evidence China wanted them to see)" (JE (CHS), June 3). Ein Beitrag macht explizit diesen Punkt, der in anderen angerissen wird oder mitschwingt: „When people understand that they cannot trust a source of information, no amount of insistence that said information comes from 'experts' or is 'based on science' will make it credible. Unfortunately, the actions of the Chinese and US governments and their media protectors will furthermore undermine our trust in the integrity in all of them" (whouck (va), June 3). Das grundlegende Misstrauen in die Daten fungiert hier als ein „undercutting defeater", der die prinzipielle Möglichkeit versperrt, den Erkenntnissen aus diesen Daten zu vertrauen.

5.4 Direkte Argumente und die Frage nach der Expertendomäne

Es sei noch auf eine letzte Ebene der Diskussion der Leserschaft des *New York Times*-Beitrages hingewiesen, die wiederum von Goldman nicht ins Auge gefasst wurde. So zeigt sich bereits in einer kursorischen Betrachtung auf frappierende Weise, wie allgegenwärtig – contra Goldman – direkte Argumente in den Kommentaren und Diskussionen sind. Direkte Argumente zielen darauf ab, eine Überzeugung durch direkten Sachbezug zu begründen. Sie erfolgen, indem Evidenz erster Ordnung präsentiert wird, etwa in Form von Erfahrungen oder logischen Gründen. Davon unterschieden sind indirekte Argumente, die auf höherstufige Evidenz rekurrieren. Letztere bestehen darin, direkte Argumente im Lichte höherstufiger Evidenz zu bewerten – also etwa das direkte Argument eines vermeintlichen Experten durch die höherstufige Evidenz zu beurteilen, dass dieser Experte einen guten Track-Record aufweist oder aber, im Gegenteil, durch einen ausgesprochenen Bias in vergangenen Stellungnahmen aufgefallen ist.

Goldman nimmt zu Recht an, dass Novizen direkte Argumente von Experten nicht bewerten können; dies erfordert nämlich Expertise, die Novizen – per Definition – nicht haben. In den Kommentaren zum *New York Times*-Beitrag führen die Lesenden jedoch eine ganze Reihe weiterer Evidenzen und (zumindest scheinbar) eigenständiger inhaltlicher, also direkter Argumente an: Es geht etwa um Evidenzen aus dem Bereich der eigenen Beobachtungen, Verweise auf Fachartikel aus dem Bereich der Virologie, um Argumente in Bezug auf realistische Beweggründe Chinas für eine mögliche Geheimhaltung der Gain-of-function-Forschung oder gar dem intentionalen Freisetzen einer Biowaffe. Auch historische Analogien zu vermeintlich vergleichbaren Vorgängen werden oft als Argumente angeführt.

Über die Überzeugungskraft der Argumente im Einzelnen wollen wir hier nicht urteilen – aufgrund ihrer Fülle wäre dies auch gar nicht sinnvoll möglich. Wir möchten aber eine Vermutung formulieren, warum gerade in Bezug auf das Beispiel solche Argumente erster Stufe besonders gut kommunikativ anschlussfähig sind. Diese Vermutung verweist auch auf eine von Goldman nicht so benannt Schwierigkeit in seinen argumentativen Setzungen.

Mit Goldman möchte man zunächst denken, dass eine solche Art der direkten Argumentation in der breiteren Öffentlichkeit eigentlich nicht wirklich naheliegend sei. Schließlich ist ihm zufolge erstens davon auszugehen, dass bei einem Artikel zu einer wissenschaftlichen Frage der Großteil der Zeitungsleserschaft keine Experten *auf diesem Gebiet* sein können. Immerhin definiert Goldman Expertise sogar darüber, dass sie durch epistemische Überlegenheit charakterisiert ist, die nur wenigen zukommt – Exklusivität ist bei ihm gewissermaßen definitorisch in den Begriff der Expertise eingebaut. Zweitens aber zeichnen sich laut Goldman Novizen dadurch aus, dass sie sich ihrer epistemischen Grenzen bewusst sind (Goldman 2001: 90). Insofern sie sich nicht als Expertinnen verstehen, wissen sie um ihre Grenzen, Argumente erster Stufe zu verstehen, geschweige denn deren Stärke einzuschätzen oder gar eigenständig ein solches Argument entwickeln zu können.

Ein Blick in das Fallbeispiel legt jedoch nahe, dass gerade diese letzte Prämisse fraglich ist: Genau genommen scheint es so zu sein, dass sich viele Personen auf der Ebene direkter Argumentation bewegen, gerade weil sie sich eben *nicht* wirklich als (völlige) Novizen verstehen. Unsere Vermutung ist, dass dies in diesem speziellen Fall gerade dadurch begünstigt ist, dass wir es mit einer Fragestellung zu tun haben, die keiner spezifischen Disziplin (*E-Domäne*, wie es bei Goldman heißt) maßgeschneidert zugeordnet werden kann. Vielmehr haben wir es mit einer sinnvollerweise multidisziplinär anzugehenden Fragestellung zu tun: Virologie, Epidemiologie, Wissenschaftssoziologie, Mikrobiologie, Bioinformatik, Biosicherheitsexpertise, Politikwissenschaft, und viele weitere können angeführt werden, um die Frage nach dem Ursprung des Virus zu verhandeln. Nicht nur Alina Chan, *niemand* kann Experte für alle diese Felder sein.[32]

Vermutlich begünstigt die *vielschichtige Multidisziplinarität in* diesem Fall, dass viele Personen in ihren Kommentaren direkte Argumente entwickeln. Auffällig ist, dass sie in diesen Fällen ihren Kommentar häufig unter Angabe ihrer Profession einleiten, die offenkundig ihren Expertenstatus hervorheben soll. Exemplarisch ist dafür die folgende Diskussion in den Kommentaren, die mit diesem Beitrag von Anne beginnt: „As a nurse who worked both the ER and ICU, I have always known it was lab grown. COVID was unlike anything other virus I came across working as a clinician. Most diseases and infections have a list of known symptoms that can generally lead a doctor to deduce you have this over that. Stuffy nose, cough, nasal drip, and headache would lead you to believe you have a head cold. Vomiting and diarrhea would indicate some for gastrointestinal issue. Covid was the first virus where you literally needed a test because it affected every single person differently. Some lost taste and smell, some simply couldn't breathe, some have had long term effects, and others had no symptoms at all, etc. How could it not be lab grown? It is too versatile of a virus affecting humans on a genetic level in ways we still are comprehending" (Anne (NJ), June 3). Daraufhin erwidert ein anderer Kommentator: „I'm an ER doctor, and I must say,

I don't follow your logic. Your conclusions don't follow from your premises" (The Golden Mean (Sydney, Australia) June 3). Und ein anderer schreibt: „With all due respect, I must confess as a physician scientist that the little hairs on the back of my neck always rise suspiciously when someone says, 'I have always known,' particularly when no data are/were available. My training includes clinical pediatrics, and I can tell you that many viral illnesses will present with symptoms that appear 'unique' to each young patient. Many viruses can cause either respiratory symptoms or gastrointestinal symptoms in affected patients; this is hardly a new phenomenon. If clinical features were able to distinguish one common viral infection from another with acceptable reliability we would not have to order so many rapid virus diagnostic tests in the ER! Sars-cov2 is hardly the first viral illness that relied upon testing for diagnosis; think how many flu tests are done each season in ERs in the US! […]" (Ivy MD (Philadelphia) June 3). Andere Kommentatoren kritisieren den Umstand, dass die Profession *Krankenschwester* keine epistemische Autorität in diesem Gebiet verleiht: „With all due respect, you're a nurse, not a virologist. There is absolutely no mystery about SARS-CoV-2's ability to affect most types of tissue, and there are other viruses with other unique abilities. Think HIV infecting T cells. Are all those viruses sure to be lab grown?" (Cornelius (Munich) June 3)

Dies sind typische Selbstverortungen, die das Anbringen von weiteren Evidenzen und Argumenten begleiten. Dass diese angebracht werden, ließe sich so deuten, dass sie gerade auf den Punkt aus unserer obigen ersten Prämisse reagieren, und dessen Aussage für sich selber ausnehmen wollen: Sie bringen zum Ausdruck, dass sie sich gerade *nicht* als vollkommene Novizen verstehen, die direkte Argumente nicht einschätzen können.

Dass nun die Fragestellung keinen klaren und singulären disziplinären Zuschnitt zulässt, begünstigt vermutlich die Anschlussfähigkeit dieses kommunikativen Zuges: Deshalb werden so viele verschiedene Disziplinen und Professionen herangezogen, um sich selbst als Experten auszuweisen (beispielsweise als Biochemiker, als Doktor der Virologie oder als Krankenschwester[33]).

Natürlich ist fragwürdig, ob diese Professionen jeweils und allgemein einschlägig genug sind, um sich sinnvoll als Experte für die spezifische Fragestellung präsentieren zu können. Die Kommentare zeigen, dass dieser Punkt auch unter den Kommentierenden umstritten sein kann. Unser Punkt ist vielmehr, dass die Prävalenz dieses Zuges darauf hinweist, dass der Disziplinenzuschnitt der Fragestellung – nicht ganz zu Unrecht – *als unscharf wahrgenommen wird.* Und genau diese Unschärfe öffnet vielfache Möglichkeiten, sich selber auch als Experten zu markieren – und stellt insofern einen nachvollziehbaren Weg zum Anbringen von Evidenz erster Ordnung dar.

6. Zur Rolle der Kriterien: Erneut zur Pervasivität von Vertrauen und Wissen

Oben haben wir bereits in der Diskussion der originär von Goldman ins Spiel gebrachten Kriterien nachgewiesen, dass diese zum Teil eine zirkuläre Struktur aufweisen: So scheint deren sinnvolle Anwendung an vielen Stellen bereits Vertrauen in andere Expertinnen voraussetzen zu müssen: Sei es Vertrauen in Experten, die einschätzen, was bedenkenswerte Interessenskonflikte oder einen bemerkenswerten Track-Record darstellen etc. Wir wollen diesen Punkt nun erweitern und nachweisen, inwiefern er auch für die am Fallbeispiel gewonnenen ergänzenden Kriterien einschlägig zu sein scheint.

Auch in dem konkreten Fallbeispiel lässt sich die Relevanz des Pervasivitätsmodells von Vertrauen und Wissen aufzeigen. Die Kriterien können nicht die Funktion erfüllen, eine vertrauensunabhängige, autarke Wissensbasis zu generieren, auf deren Grundlage man das Vertrauen und Misstrauen in Experten orientieren kann. Vielmehr soll auch hier das pervasive Verhältnis von Vertrauen und Wissen (in der Form der Vertrauensgeladenheit oder des eingekapselten Vertrauens) freigelegt werden.

Am naheliegendsten ist dies vermutlich für das oben genannte Kriterium des medialen Kontextes (5.2) eines Expertenauftritts.

Dieser funktioniert nur dann als Orientierungsgröße für Expertenvertrauen, wenn bereits Vertrauen in die entsprechenden medialen Hintergrundinstitutionen als Voraussetzung besteht. Wenn man diesen Punkt weiterdenkt, wird der Sachverhalt aber schnell verwickelt: Worauf wiederum könnte sich ein solches Vertrauen vernünftigerweise stützen? Woher meinen wir eigentlich zu wissen, dass bestimmte Medieninstitutionen durch eine vernünftige journalistische und redaktionelle Praxis getragen sind?

Auch hier zeigt sich schnell, dass sich dies schwerlich auf eine reine Wissensbasis zurückführen lässt. Wann und in welcher Form kommen wir etwa in direkten, augenscheinlichen Kontakt mit solchen institutionalisierten Praktiken? Selbst wenn man sich etwa vorstellte, man schaue sich eine filmische Dokumentation über den journalistischen Alltag einer Redaktion an: Ich brauche bereits ein Hintergrundwissen darüber, dass es regelbasierte Praktiken und institutionelle Arrangements gibt, die in dem, was ich sehe, verkörpert sind (Coadys Punkt). Ich muss zudem darauf vertrauen, dass das, was ich sehe, nicht gestellt ist, sondern eine praxisnahe Wirklichkeit abbildet, damit ich zu Wissen über die abgebildeten journalistischen Tätigkeiten kommen kann (Lahnos Punkt). Auch die Schulung von Medienkompetenz, die verspricht, dass vertrauenswürdige Medien besser erkannt werden können, wird ein entsprechendes Grundvertrauen voraussetzen müssen.

Ebenso, wenn auch subtiler, setzt die Anwendung des Kriteriums der Motive bzw. Tugenden (5.1) Vertrauen voraus. Die Idee mit Blick auf dieses Kriterium war, dass man an dem Verhalten der Akteure direkt und voraussetzungslos ablesen könne, ob es sich dabei um epistemisch tugendhafte Personen handelt oder nicht – etwa, da sie eine offene und selbstkritische Geisteshaltung zeigen und sich als nicht zu engstirnig gegenüber anderen Positionen erweisen. Die Schwierigkeit dabei scheint jedoch zu sein, dass sich gar nicht so ohne Weiteres beurteilen lässt, was tugendhaft unvoreingenommene und lasterhaft engstirnige Einstellungen sind – *ohne, dass man bereits etwas darüber voraussetzt, was evidenziell begründete und ernstzunehmende Positionen sind, und was nicht*.

Epistemische Tugenden haben es wie alle Tugenden an sich, dass sie kontextsensitiv sind: Die Tugend der Freundlichkeit zu besitzen bedeutet etwa, nur dann und in dem Maße freundlich zu sein, als es in einer konkreten Situation angemessen ist – es eine andere Person etwa verdient hat, dass man ihr gegenüber Freundlichkeit zeigt. Für epistemische Tugenden wie intellektuelle Offenheit gilt aber ebenso: Als Tugenden kennen sie auch ihre angemessenen Grenzen. Die Grenzen liegen in diesem Falle darin, was aus wissenschaftlicher Sicht ernst zu nehmen ist. Betrachten wir etwa eine Klimawissenschaftlerin, die klimaskeptische Positionen nicht ernst nimmt, da diese auf bekannten und bereits widerlegten Argumenten beruht. In diesem Falle spräche es nicht gegen ihre epistemische Tugendhaftigkeit, wenn sie sich von vornherein nicht auf eine Diskussion mit solchen Positionen einlässt – vermutlich eher im Gegenteil.

Epistemische Tugendhaftigkeit lässt sich also nicht unabhängig von einer vorausgesetzten Einschätzung der epistemischen Situation erkennen.[34] Erneut führt dies aber auf die eigentliche Ausgangsfrage zurück: Wie können Novizen denn die epistemische Situation angemessen einschätzen – ohne bereits gewissen Expertinnen vertrauen zu müssen? Die vermeintlich direkte Wahrnehmung des epistemisch tugendhaften/lasterhaften Auftretens der Expertinnen lässt sich somit wiederum als ein von Coady beschriebener Fall deuten: Das Kriterium kann nicht auf bloßem *Augenschein* beruhen; und daher kann Vertrauen nicht auf *vertrauensfreie* Evidenz zurückgeführt werden. Vielmehr ist die evidenzielle Ebene bereits vertrauensgeladen: Um das kommunikative Verhalten der Experten *als* Ausdruck eines epistemischen Lasters oder einer epistemischen Tugend zu erkennen, sind Novizen auf eine Einschätzung der Evidenzlage als interpretative Hintergrundfolie angewiesen. Diese wiederum, analog zum Queen-Beispiel von Coady, kann selber nicht auf rein autarker Wissensbasis gewonnen werden.

Zuletzt setzen auch die Kriterien *Herkunft der Daten* (5.3) und *Expertendomäne* (5.4) bereits vorstrukturierende Vertrauens- bzw. Misstrauensbeziehungen voraus: Für 5.3 ist dies eben

Misstrauen in die Daten aus China, das sich somit wiederum in Misstrauen gegenüber den Ergebnissen der Untersuchung fortpflanzt. Für den Punkt von 5.4 wiederum muss bereits ein Verständnis vorausgesetzt werden, welche Disziplinen eigentlich zur Beantwortung der Frage nach dem Virusursprung einschlägig sind. Dafür wiederum scheint ein höherstufiges Vertrauen vorausgesetzt werden zu müssen: nicht nur in die Positionen oder jeweiligen Antworten, sondern mit Blick auf die Aussage, wer überhaupt – welche Profession – eine angemessene Perspektive zur Beschreibung und Beantwortung des Problems liefern kann.

7. Fazit

Es wurde schon des Öfteren darauf hingewiesen, wie zentral eigentlich Vertrauen für unser Wissen, aber auch generell für unsere epistemischen Überzeugungen, Annahmen, Vermutungen etc. ist. Dies ist uns größtenteils nicht bewusst, zuweilen gibt es gar eine Blindheit gegenüber der enormen Relevanz, die Vertrauen für unsere Erkenntnisprozesse eigentlich spielt.

Das hier vorgeschlagene Pervasivitätsmodell geht darüber hinaus aber noch einen Schritt weiter, indem es auf die mehrfachen *Verstrickungen* von Vertrauen und Wissen aufmerksam macht und auf die Unmöglichkeit in vielen Fragen eine vertrauensunabhängige Wissensbasis zu erlangen. Vertrauen ist demnach nicht nur zentral, es scheint überhaupt keinen epistemischen Bereich zu geben, der ohne Vertrauen auskommt. Der größte Teil unseres Wissens, wenn nicht sogar all unser Wissen, scheint bereits vertrauensgeladen zu sein oder Vertrauen einzukapseln.

Das Pervasivitätsmodell von Vertrauen und Wissen erwies sich zudem als ein fruchtbarer begrifflicher Rahmen, um über das Novize-2-Expertenproblem nachzudenken. Vor diesem Hintergrund haben wir den verschiedenen Verstrickungen von Vertrauen und Wissen in unserem Fallbeispiel nachgespürt und diese freigelegt. Auch hier zeigt sich, dass die vielfach eingebrachten Kriterien nicht auf vertrauensfreies Wissen zurückführbar sind –

oft verdeckt hinter einer gegenteiligen Rhetorik der beteiligten Akteure, die eher den Eindruck erwecken, sich autark und evidenzbasiert auf bloßes Wissen zu beziehen. Hingegen: selbst etwas vermeintlich so Augenscheinliches wie die epistemische Tugendhaftigkeit von Akteuren scheint sich nicht ohne vertrauensbasiertes Wissen einschätzen zu lassen.

Anmerkungen

[1] Die Philosophie von Vertrauen, die Philosophie der Zeugenschaft sowie die soziale Epistemologie sind nicht identisch, so muss sich z.B. Vertrauen nicht primär auf Aussagen (Bezeugungen) beziehen; gleichwohl gibt es eine große Überschneidung nicht nur mit Blick auf den Gegenstand, sondern auch hinsichtlich der jeweils entwickelten Theorien.

[2] Diese Einwände gehen über den zentralen Einwand hinaus, dass den epistemischen Ansätzen ein Kategorienfehler unterläuft, da Vertrauen bei ihnen auf Formen der Prognose reduziert wird. Dieser Einwand besagt, dass man Vertrauen nicht angemessen begreifen kann, wenn man es rein oder primär als eine epistemische Beziehungsform denkt.

[3] Wir betrachten im Folgenden primär den letzten Fall. Im Grunde ist das Problem, um das es Goldman geht, für konträre oder widersprechende Urteile das gleiche (mit Blick auf die Schwierigkeiten von Novizen, sich zu diesen Urteilen zu verhalten). Widersprechende Urteile scheinen aber aus heuristischen Gründen eine hilfreiche Zuspitzung des Problems zu liefern.

[4] Es gibt Fälle, die klarer erscheinen mögen. Wenn etwa eine vermeintliche Expertin in offensichtlicher Weise selbst für Novizen unsinnige Behauptungen macht. Dieser Fall wäre aber dann einer, in der eine direkte Bewertung für den Novizen zugänglich ist; es wäre damit ein Fall, den wir hier gerade nicht betrachten und der ungleich seltener gegeben ist. Zudem wäre in diesem Fall unklar, ob es sich überhaupt um eine Expertin handelte.

[5] Der von Goldman vorgeschlagene Kriterienkatalog, wurde bereits mehrfach diskutiert und kritisiert – zuweilen als ergänzungsbedürftig, als in Teilen nicht praktikabel oder in Teilen nicht ausreichend reliabel. Während ein Großteil dieser Diskussion eher den genauen

Zuschnitt des Kataloges in den Blick nahm, hat Neil Levy (2024) eine grundlegendere Kritik formuliert, der der unseren hier strukturell ähnlich ist: Levy zeigt mit Blick auf Goldmans Kriterien, dass diese entweder bereits Expertise voraussetzen müssen, oder, insofern dies von Novizen nicht angenommen werden kann, sie in Teilen bereits bestehendes Vertrauen in andere Instanzen stillschweigend voraussetzen. Am einfachsten ist dies vermutlich für das Kriterium der Meta-Experten einsehbar. Dieses setzt offensichtlich bereits Vertrauen in entsprechende Meta-Experten voraus. Entsprechend kann eine Person, die dieses Vertrauen nicht hat, nicht ohne Weiteres vom Expertenstatus einer Person überzeugt werden. Umgekehrt: Wenn jemand Vertrauen in *alternative* (vermeintliche) Meta-Experten hat, wird sie mit den *gleichen* Kriterien zu ganz *unterschiedlichen* Einschätzungen gelangen. Wir ergänzen in diesem Abschnitt Levys Überlegungen mit einigen eigenen Argumenten. Über Levy hinausgehend wird unser Ziel im weiteren Verlauf sein, dieses Problem als eines der verschiedenartigen Verwickeltheit von Vertrauen und Wissen zu artikulieren.

[6] Eigentlich in zwölf Fällen, der letzte ist jedoch eher summierend und nennt ein generelles Problem von Streecks Einschätzungen.

[7] Ausgehend von diesen und weiteren Überlegungen entwickelt Coady seine Theorie einer kohärenz- und kohäsionsbasierten Epistemologie der Zeugenschaft, in der es ferner apriorische Gründe für die Vertrauenswürdigkeit anderer gibt (Coady 1992: 152–172). Diese weitergehende Theorie spielt für unsere Argumentation jedoch keine Rolle.

[8] „A wise man, therefore, proportions his belief to the evidence." (Hume 2007 [1748]: 80)

[9] Coady macht in seiner Kritik an Humes Vorschlag, wie die Vertrauenswürdigkeit von Zeugen zu bewerten ist, einen ähnlichen Punkt. So wendet er gegen Hume ein, dass dessen Versuch, Vertrauen auf ein von Vertrauen unabhängiges Wissen zu gründen, unter anderem daran scheitert, dass in den meisten Fällen dieses Wissen nicht selbst gewonnen werden kann (Coady 1992: 80–82). Betrachten wir etwa folgende Situation: Wir wollen bewerten, wie vertrauenswürdig ein Zeuge ist. Humes Verfahren gemäß würden wir prüfen, wie viele der Berichte, die er gegeben hat, wahr waren. Dies würde uns auf ein Verhältnis (Anzahl wahrer Berichte/Anzahl seiner Berichte insgesamt) bringen. Dieses Verfahren setzt also voraus, dass wir die

Wahrheit der jeweiligen Berichte kennen. Aber – und das ist für Coady ein entscheidender Punkt – nur in wenigen Fällen *wissen* wir, ob der jeweilige Bericht wahr war aufgrund eigener Erfahrung oder eigenen logischen Nachdenkens. Nur in wenigen Fällen wissen wir dies autark. Denken wir an journalistische Reportagen, an Nachrichten in sozialen Netzwerken, an wissenschaftliche Studien, so gilt fast ausnahmslos, dass wir die dort behaupteten Dinge nicht selbst erfahren haben. Und selbst bei dem, was uns ein Nachbar über einen anderen erzählt, sind wir vermutlich nicht Zeuge gewesen – weshalb er es uns ja mitteilt. Wenn das der Fall ist, dann können wir aber fast nie die Berichte anderer auf der Grundlage dieses Verfahrens prüfen; und wir können fast nie einen gut begründeten Eindruck von ihrer Vertrauenswürdigkeit gewinnen. Denn letzteres setzte voraus, dass wir nicht nur in einem Fall, sondern in vielen Fällen ‚wissen', ob ihre Berichte wahr waren oder aber nicht. Nun bewerten wir jedoch die Glaubwürdigkeit dieser Berichte nahezu pausenlos. Und das liegt daran, dass wir dieses Wissen von der Wahrheit der Berichte anderer auf der Grundlage gewinnen, dass uns andere mitteilen, ob ein Bericht wahr ist oder nicht. Das heißt: Wir bewerten, wie Coady verdeutlicht, Bezeugungen und Zeugen anhand der Bezeugungen *anderer* Zeugen. Wir stoßen damit erneut auf den Gedanken, dass dieses Wissen nicht autark ist, sondern vom Vertrauen in andere abhängt. Damit kommt es aber zu einem zirkulären Verhältnis von Vertrauen und Wissen: E_1 berichtet p. E_2 versichert, dass p falsch ist. N bewertet den Bericht von $E_{1,}$ dass p auf der Grundlage des Berichts von E_2, dass p falsch ist. N mindert daher die Vertrauenswürdigkeit von E_1 auf der Grundlage von seinem Vertrauen in E_2. Diese abstrakte Überlegung wird von uns alltäglich mit großer Selbstverständlichkeit vollzogen: Wenn Katja etwa einen Politiker für unglaubwürdig hält, dann kommt sie zu diesem Schluss, da sie den Berichten anderer vertraut, die den Behauptungen des Politikers widersprechen. Wenn Mira der Bewertung eines Produkts misstraut, dann tut sie dies auf der Grundlage anderer Bewertungen des gleichen Produkts, denen sie vertraut und die mit der ersten Bewertung nicht in Einklang zu bringen sind. Oder: Wenn wir glauben, dass Streecks Prognosen einen schlechten Track-Record aufweisen, dann kommen wir zu dieser Bewertung, indem wir seine Prognosen etwa als unzutreffend bewerten – auf der Grundlage der Berichte anderer, denen wir vertrauen (etwa mit Blick darauf, wie hoch die Sterblichkeit an

Covid-19 ist). Auch hier sehen wir also: In solchen Fällen ist Evidenz *nicht unabhängig* von Vertrauen. Evidenz wird vielmehr zu Evidenz für etwas, wenn wir Quellen dieser Evidenz vertrauen. Anders gesagt: Evidenz spricht hier nicht für sich selbst, sondern nur indem andere sprechen, denen wir vertrauen.

[10] Vgl. dazu Andersen u.a. 2020.

[11] Faucis genaue Worte waren: „There was a study recently that we can make available to you, where a group of highly qualified evolutionary virologists looked at the sequences there and the sequences in bats as they evolve. And the mutations that it took to get to the point where it is now is totally consistent with a jump of a species from an animal to a human. So, I mean, the paper will be available — I don't have the authors right now, but we can make that available to you." Das Transkript der PK ist abrufbar: https://trumpwhitehouse.archives.gov/briefings-statements/remarks-president-trump-vice-president-pence-members-coronavirus-task-force-press-briefing-april-17-2020/ (letzter Aufruf 05.01.2025). Spätere Mails legten nahe, dass Fauci die Autoren rund um den „Proximal Origin"-Artikel zusammenführte und z.T. auch mit ihnen in einer Diskussion darüber war. Unter anderem deshalb wurde die Weise, in der er auf das Paper verwies, als gezieltes Geheimhalten und daher Täuschung von einigen Personen ausgelegt. Er selbst verwies darauf, dass er die Namen, während der Pressekonferenz nicht erinnerte und auch nicht am Schreibprozess beteiligt war. Vgl. dazu auch den BBC-Podcast *The Hunt for Covid's Origin: Episode 3. Uncertainty and Trump*, Jun 13, 2023.

[12] Vgl. den Bericht der BBC zur Pressekonferenz. Hier abrufbar: https://www.bbc.com/news/world-us-canada-52496098 (zuletzt aufgerufen am 08.01.2025).

[13] Vgl. hierzu Seite 9 des Berichts, abrufbar: https://www.who.int/publications/i/item/who-convened-global-study-of-origins-of-sars-cov-2-china-part (zuletzt aufgerufen am 07.01.2025).

[14] https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30418-9/fulltext (zuletzt aufgerufen am 17.12.2024).

[15] https://www.science.org/doi/10.1126/science.abj0016 (zuletzt aufgerufen am 17.12.2024).

[16] Auch innerhalb der WHO wurde im Sommer 2021 der unzureichende Zugang zu „raw data" kritisiert und weiterer Untersuchungsbedarf gesehen. Vgl. die Äußerungen des WHO Generaldirektors

Tedros Adhanom Ghebreyesus, hier dokumentiert: https://web.archive.org/web/20210731031020/https://www.reuters.com/article/us-health-coronavirus-who-china-report-idUSKBN2BM26S (zuletzt aufgerufen am 07.01.2025).

[17] Dabei handelt es sich um die *D*efense *A*dvanced *R*esearch *P*rojects *A*gency, es ist eine Behörde des Verteidigungsministeriums der Vereinigten Staaten, die Forschungsprojekte durchführt.

[18] Geleaked wurde Antrag von DRASTIC Research, einem Kollektiv anonymer Wissenschaftlerinnen, die sich mit Bioforschung und Biosafety befassen. Vgl. hierzu: https://drasticresearch.org (zuletzt aufgerufen am 29.12.2024).

[19] Vgl. hierzu die Kommentierung durch die Nonprofit Organisation *U.S. Right to Know*, die auch die Mails publizierte. „Daszak noted that he could 'send it round' to other scientists to sign. 'We'll then put it out in a way that doesn't link it back to our collaboration so we maximize an independent voice,' he wrote. The two scientists Daszak wrote to about the need to make the paper appear independent of EcoHealth, are coronavirus experts Ralph Baric and Linfa Wang. In the emails, Baric agreed with Daszak's suggestion not to sign The Lancet statement, writing 'Otherwise it looks self-serving, and we lose impact.' Daszak did ultimately sign the statement himself, but he was not identified as its lead author or coordinator of the effort." Zu finden: https://usrtk.org/covid-19-origins/scientists-masked-involvement-in-lancet-letter-on-covid-origin/# (zuletzt aufgerufen am 07.01.2025).

[20] Wir lassen hier etwa die gesamte Geschichte der US-Geheimdienste und ihrer sich teilweise wandelnden Einschätzungen aus; ferner die Informationen, die ihm Rahmen verschiedener Kommissionen des US-Kongress veröffentlicht wurden.

[21] Vgl. etwa Holmes, Edward C. et al. 2021. Vgl. Auch die Expertenumfrage vom Februar 2024, hier zu finden: https://gcrinstitute.org/papers/069_covid-origin.pdf (zuletzt aufgerufen am 14.01.2025). In der Umfrage spricht sich eine sehr große Mehrheit der befragten Expertinnen dafür aus, dass das Virus natürlich-zoonotischen Ursprungs ist. Das setzt natürlich weiterhin voraus, dass Novizen vertrauen, dass Expertinnen befragt wurden und man diesen Expertinnen zutraut, zu den richtigen Schlüssen zu kommen.

[22] Vgl. den Bericht der *Zeit* online vom 19.03.2025 „Nicht nur der Geheimdienst wird langsam stutzig", in dem ohne die Evidenz selbst

vorzustellen, berichtet wird, dass der BND frühzeitig die Laborhypothese für deutlich wahrscheinlicher hielt und dass mehrere Bundesregierungen sich entschieden, dies geheim zu halten. Darin wird auch die Bewertung der Evidenz durch Wissenschaftler geschildert. https://www.zeit.de/gesundheit/2025-03/bnd-bericht-corona-ursprung-labor-pandemie (letzter Aufruf 31.10.2025).

[23] Vgl. Pekar, Jonathan E. et al. 2022.

[24] Ebenso nennt Anderson „[c]onflicts of interest, such as receiving funds from agents who have stake in getting people to believe a particular claim" als starkes Motiv für Experten, die Unwahrheit zu sagen (Anderson 2011: 147).

[25] Zur Diskussion um die angemessene Konzeption solcher epistemischen Tugenden siehe etwa Fairweather/Zagzebski 2001.

[26] Zum generellen Unterschied zwischen bloßen Fertigkeiten und Tugenden siehe etwa Halbig 2013: 73–77.

[27] Dass dies bei Goldman selber etwa nicht vorkommt, kann viele Gründe haben. Vermutlich liegt einer darin, dass Goldman paradigmatisch eine Situation vor Augen hat, in der zwei Experten im gleichen Medium aufeinandertreffen, etwa im Format eines übertragenen Streitgesprächs. Zudem ist davon auszugehen, dass in der Zeit seines Aufsatzes im amerikanischen Kontext, die Medienlandschaft weniger polarisiert war und sich vor allen Dingen Streitpositionen in Bezug auf Wissenschaften weniger klar verschiedenen Medienlagern zuordnen ließen.

[28] Auch Elizabeth Anderson erwähnt Peer-Review, jedoch nicht auf die mediale Dimension, sondern wiederum indirekt auf die charakterliche Dimension bezogen. Das Umgehen des Peer-Review-Verfahrens könne nämlich als Ausdruck fehlender epistemischer Verantwortlichkeit verstanden werden (Anderson 2011: 147).

[29] Allerdings kann tatsächlich gezweifelt werden, ob ein Artikel wie der von Chan überhaupt gut in einer wissenschaftlichen Fachzeitschrift untergebracht werden kann. Schwierigkeiten kann hier bereiten, dass die vorgebrachten Argumente nicht einer eindeutigen Fachdisziplin zugehörig scheinen (dazu unten mehr). Zudem ist zu bedenken, dass es der Autorin vermutlich gerade darum geht, eine breitere Öffentlichkeit direkt zu adressieren – wofür ein *New York Times*-Meinungsbeitrag wohl schlichtweg effektiver ist als eine Fachpublikation.

[30] Er berichtete, dass er mitten an der Arbeit am Podcast schnellstmöglich mit seiner Familie China verlassen musste, da er schwerwiegende Sanktionen zu erleiden befürchtete.

[31] Vgl. auch eine andere Passage im Podcast, in der es um den Besuch des China-Teams, das den Ursprung von Covid-19 im Auftrag der WHO untersucht, im WIV geht. Angesprochen auf den Punkt, dass sie keinen Zugang zu den Rohdaten der Blutanalysen der Mitarbeiterinnen des WIV erhielten, die hätten zeigen können, ob diese Antikörper im fraglichen Zeitraum hatten, was auf eine Coronainfektion hindeuten würde, antwortete Marion Koopmans dem Journalisten Sudworth: „We have asked questions we felt even we discussed it among ourselves. Wow, that was tough. I mean, are you sure? How did you check? Did you test your people? But it was good. We all felt they needed to be asked, but we also *felt* that the answers were *credible*." (The Hunt for Covid's Origin, Episode 4, ab 24:03). Dass Koopmans davon spricht, dass es sich anfühlte, als seien die Antworten glaubwürdig, bedeutet nicht, dass es keine epistemischen Gründe für sie gab, aber dass Vertrauen bei diesen wohl unweigerlich eine Rolle spielte.

[32] Dies ist eine der Schwierigkeiten, problematisches „trespassing" von unproblematischen Fällen des fachlichen Grenzübertritts abzugrenzen. Es ist ein Problem, das sich nicht nur in interdisziplinären Forschungszusammenhängen zeigt, sondern bei gesellschaftlichen Fragen, die vielfältige Aspekte aufweisen und somit mannigfache Perspektiven zulassen. (Vgl. Ballantyne 2019; DiPaolo 2022)

[33] Siehe etwa die Kommentare von: Anne (NJ); The Golden Mean (Sydney, Australia); Dan (Gainesville, Florida); MEK (Berkeley) etc.

[34] Vgl. Coady/Corry 2013: 29f. für ein ähnliches Argument in Bezug auf vermeintliche Biases.

Literatur

Andersen, Kristian G., Andrew Rambaut, W. Ian Lipkin, Edward C. Holmes und Robert F. Garry (2020) „The Proximal Origin of SARS-CoV-2". *Nature Medicine* 26 (4): 450–452. https://doi.org/10.1038/s41591-020-0820-9.

Anderson, Elizabeth (2011) „Democracy, Public Policy, and Lay Assessments of Scientific Testimony". *Episteme* 8 (2): 144–164. https://doi.org/10.3366/epi.2011.0013.

Ballantyne, Nathan (2019) „Epistemic Trespassing". *Mind* 128 (510): 367–395. https://doi.org/10.1093/mind/fzx042.

Bloom, Jesse D., Yujia Alina Chan, Ralph S. Baric, u.a. (2021) „Investigate the origins of COVID-19". *Science* 372 (6543): 694–694. https://doi.org/10.1126/science.abj0016.

Calisher, Charles, Dennis Carroll, Rita Colwell, u. a. (2020) „Statement in Support of the Scientists, Public Health Professionals, and Medical Professionals of China Combatting COVID-19". *The Lancet* 395 (10226): e42–43. https://doi.org/10.1016/S0140-6736(20)30418-9.

Chan, Alina (2024) „Opinion | Why the Pandemic Probably Started in a Lab, in 5 Key Points". *The New York Times* (3. Juni 2024), Abschn. Opinion. https://www.nytimes.com/interactive/2024/06/03/opinion/covid-lab-leak.html.

Coady, C.A.J. (1992) *Testimony: A Philosophical Study*. Oxford; New York: Oxford University Press. https://doi.org/10.1093/0198235518.001.0001.

Coady, David und Richard Corry (2013) *The Climate Change Debate. An Epistemic and Ethical Enquiry*. Basingstoke, Hampshire: Palgrave Mcmillan.

DiPaolo, Joshua (2022) „What's Wrong with Epistemic Trespassing?" *Philosophical Studies* 179 (1): 22343. https://doi.org/10.1007/s11098-021-01657-6.

Fairweather, Abrol und Linda Zagzebski (Hrsg.) (2001) *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*. Oxford University Press.

Goldman, Alvin I. (2001) „Experts: Which Ones Should You Trust?" *Philosophy and Phenomenological Research* 63 (1): 85–110. https://doi.org/10.2307/3071090.

Halbig, Christoph (2013) *Der Begriff der Tugend und die Grenzen der Tugendethik*. Berlin: Suhrkamp.

Hardwig, John (1991) „The Role of Trust in Knowledge". *The Journal of Philosophy* 88 (12): 693–708. https://doi.org/10.2307/2027007.

Hibbett, Maia und Sharon Lerner (2021) „Leaked Grant Proposal Details High-Risk Coronavirus Research". *The Intercept* (23. September 2021). https://theintercept.com/2021/09/23/coronavirus-research-grant-darpa/.

Holmes, Edward C., Stephen A. Goldstein, Angela L. Rasmussen, David L. Robertson, Alexander Crits-Christoph, Joel O. Wertheim, Simon J. Anthony, u.a. (2021) „The Origins of SARS-CoV-2: A Critical Review". *Cell* 184 (19) (16. September 2021): 4848–4856. https://doi.org/10.1016/j.cell.2021.08.017.

Hume, David (2007 [1748]) *An Enquiry Concerning Human Understanding*. Oxford: Oxford University Press.

Kaminski, Andreas (2024) „Trust in Science. A Systematic-Historical Introduction". In: Michael M. Resch, Nico Formánek, Ammu Joshy, Andreas Kaminski (Hrsg.): *The Science and Art of Simulation: Trust in Science*. Cham: Springer Nature, 1–28.

Lahno, Bernd (2002) *Der Begriff des Vertrauens*. Paderborn: Mentis.

Levy, Neil (2024) „No Trespassing! Abandoning the Novice/Expert Problem". *Erkenntnis* 90 (1): 2077–2094. https://doi.org/10.1007/s10670-024-00794-8.

Mößner, Nicola (2018) „Trusting the Media? TV News as a Source of Knowledge". *International Journal of Philosophical Studies* 26 (2): 205–220. https://doi.org/10.1080/09672559.2018.1450079.

Origgi, Gloria (2022) „The Social Indicators of the Reputation of an Expert". *Social Epistemology* 36 (5): 541–549. https://doi.org/10.1080/02691728.2022.2116962.

Pekar, Jonathan E. u.a. (2022) „The Molecular Epidemiology of Multiple Zoonotic Origins of SARS-CoV-2". *Science* 377 (6609): 960–966. https://doi.org/10.1126/science.adl0585.

Scholz, Oliver (2009) „Experts: What They Are and How We Recognize Them – a Discussion of Alvin Goldman's View". *Grazer Philosophische Studien* 79 (1): 187–205. https://doi.org/10.1163/18756735-90000864.

Schwägerl, Christian und Joachim Budde (2021) „Der Mann der dauernd falsch liegt, aber immer wieder als Corona-Experte gebucht wird". *Übermedien* (9. Februar 2021). https://uebermedien.de/57343/hendrik-streeck-der-mann-der-dauernd-falsch-liegt-aber-immer-wieder-als-corona-experte-gebucht-wird/.

Sellars, Wilfrid (1963) „Empiricism and the Philosophy of Mind". In *Science, Perception and Reality*, von Wilfrid Sellars, Atascadero, Calif: Ridgeview Pub, 127–196.

Sudworth, John (2023) „A Question of Trust – BBC Sounds". *BBC Radio 4 – Fever: The Hunt for Covid's Origin*. https://www.bbc.co.uk/sounds/play/m001mlmy.

KAPITEL 8

Vom Zählen zum Bewerten – das Wunder objektiver Wissenschaftler-Beobachtung

Klaus Erlach

Primäre Aufgabe der Wissenschaft ist es, wie der Name schon sagt, Wissen zu schaffen, also neues Wissen zu generieren und bereitzustellen. Dieses Wissen soll unterschiedliche Kriterien der Wissenschaftlichkeit erfüllen wie beispielsweise Tatsachenkonformität und Widerspruchsfreiheit, um im wissenschaftlichen Diskurs als gültig anerkannt zu werden. Wie die über ein Jahrhundert andauernde Diskussion der Wissenschaftstheorie gezeigt hat, ist der Nachweis der Gültigkeit wissenschaftlicher Aussagen alles andere als trivial. Hinzu kommt, dass wissenschaftshistorische (Fleck, Kuhn) und wissenschaftssoziologische (Rheinberger, Knorr-Cetina) Untersuchungen gezeigt haben, dass die wissenschaftliche Praxis von diesen theoretischen Überlegungen doch sehr stark abweicht – und dabei trotzdem als wissenschaftlich korrekt angesehen wird.

Zitationsvorschlag für dieses Kapitel:
Erlach, K. 2026. Vom Zählen zum Bewerten – das Wunder objektiver Wissenschaftler-Beobachtung. In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* S. 223–246. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.j. License: CC BY 4.0

Die metatheoretische Unklarheit darüber, was definitiv als wissenschaftlich anzuerkennen ist und was nicht, kann weitreichende Auswirkungen auf die Legitimierung von Wissenschaft überhaupt haben. Dies zeigt sich derzeit auch an einer teilweise wissenschaftskritischen bis wissenschaftsfeindlichen Debatte, was hier jedoch nicht das Thema sein soll. Eine innerwissenschaftliche Auswirkung der infragestehenden Legitimierung ist, dass man sich auf sehr heterogen verfasste Disziplinen und plurale methodische Zugänge einstellen muss. Diese lassen sich nicht (mehr) zu einem hierarchisch aufgebauten, konsistenten und zeitlos gültigen Wissenschaftssystem konsolidieren, wie es vielleicht noch Hegel vorschwebte, sondern sie entsprechen eher einem zeitlich dynamischen Netzwerk von unterschiedlichen, jeweils disziplinär verfassten ‚Wissensansichten'. Für den gleichen Untersuchungsgegenstand – z.B. ‚der Mensch' – gibt es dann dementsprechend ganz unterschiedliches disziplinäres Wissen – sei es evolutionsbiologisch, genetisch, physiologisch, medizinisch, psychologisch, kognitionswissenschaftlich, anthropologisch, paläontologisch, ethnologisch, ökonomisch, ergonomisch (diese Liste ist nicht auf Vollständigkeit überprüft).

Diese heterogenen Wissensansichten können dabei nach jeweils eigenen Regeln der Fachwissenschaft und das ggf. nur innerhalb eines spezifischen Paradigmas als gültiges Wissen anerkannt werden. Auf der Arbeitsebene der Wissenschaft ist dies aber keine bloße Ja-Nein-Entscheidung wie bei nobelpreisverdächtigen Fundamentalgesetzen. Vielmehr stellt sich auch nach prinzipieller Zustimmung noch immer die Frage nach der relativen Güte und Originalität einzelner Arbeiten, also die Frage nach der *Qualität* der wissenschaftlichen Ergebnisse und davon abgeleitet aus wissenschaftspragmatischen Gründen die Frage nach der *Exzellenz* der jeweiligen Wissenschaftler. Da nun offenbar kein allgemeingültiger, objektiver Bewertungsmaßstab zur Verfügung steht, zeigt sich das Legitimierungsproblem der Wissenschaft auch auf der Ebene der internen Wissenschaftssteuerung. Traditionell ‚gelöst' wurde dieses Problem durch das Peer-Review-Verfahren, bei dem Experten der gleichen Disziplin

mit ihren jeweils gültigen Maßstäben ihre Kollegen verfahrensbedingt tendenziell subjektiv bewerten.

Die ‚Wissensexplosion' der letzten Jahre, erkennbar an der steil steigenden Publikationszahl, hat eine Bewertung von Wissen als Orientierungshilfe zusätzlich unverzichtbar gemacht und gleichzeitig das Review-Verfahren an die Grenze seiner Leistungsfähigkeit gebracht. So haben sich die bibliometrischen Verfahren, ursprünglich für die Literaturauswahl in Bibliotheken gedacht, mittlerweile durchgesetzt (vgl. dazu den Beitrag „Peer Review, Bibliometrie, Altmetrics" von B. Mittermaier in diesem Band), weil sie gerade zur Hand waren und scheinbar den ersehnten objektiven Bewertungsmaßstab mitbringen, weil sie auch bei großen Informationsmengen automatisiert ermittelbar sind und weil sich das – auch nicht ganz unwichtig – hervorragend kommerziell abbilden, um nicht zu sagen: ausnutzen, lässt (Holzer 2022: 165 f.). Dass dieses scheinbar objektive Vorgehen zahlreiche Mängel hat, ist vielfach gezeigt worden (z.B. Retzlaff 2022). Dies aufgreifend soll im Folgenden ausgehend von der Frage, wie welche Wissensform grundsätzlich bewertet werden kann (Abschnitt 1), kurz dargestellt werden, welche zentralen Mängel die aktuelle, scheinbar objektive Wissenschaftsbewertung aufweist (Abschnitt 2). Da es nicht zielführend erscheint, diese Mängel durch komplette Reduktion des Wissens auf Zähl- und Messgrößen wegobjektivieren zu wollen, soll daran anschließend das ganze Spektrum möglicher Bewertungsverfahren von stark subjektiv bis absolut objektiv aufgezeigt werden (Abschnitt 3). In einem Exkurs werden ergänzend die Methoden der Qualitätssicherung in der Industrie in ihrer historischen Entwicklung betrachtet. Wer hier primär objektivere Bewertungsverfahren als in der Wissenschaft erwartet, wird enttäuscht: Das Qualitätswesen in der Industrie hat sich vom reinen Zählen und Messen der Produktionsergebnisse hin zur Verbesserung der Produktionsprozesse selbst im Qualitätsmanagement entwickelt (Abschnitt 4). Ein Lösungsansatz könnte also sein, Methoden des Qualitätsmanagements für die Wissenschaft (und nicht nur in der Lehre) zu entwickeln (Abschnitt 5). Der Beitrag schließt mit kurzen

Überlegungen dazu, was dies im Fach ‚Philosophie' bedeuten könnte (Abschnitt 6).

1. Drei Ebenen der Wissenspyramide zur Wissenschaftsbewertung

Angeregt von T.S. Eliots Gedicht *The Rock* hat Robert Lucky das Modell der *Wissenspyramide* entwickelt (Lucky 1989: 20), das sich auf Grund seiner anschaulichen Begriffsstrukturierung in der Informationswissenschaft weit verbreitet hat. Mit den Zeilen „Where is the wisdom we have lost in knowledge? / Where is the knowledge we have lost in information?" (Eliot 1934: 7) wird nicht nur eine dreistufige Hierarchie aufgespannt, sondern es werden zugleich auch die Verluste benannt, die beim Abstieg über die Ebenen entstehen. Umgekehrt heißt das, dass beim Aufstieg immer etwas hinzukommen muss, was in der jeweils unteren Ebene nicht bereits enthalten ist.

Das Basismodell der Wissenspyramide unterscheidet die drei hierarchischen Ebenen Daten – Information – Wissen (Aarmodt 1995: 198). Mit Daten sind Zeichen (genauer: Zeichenkörper) in korrekter Syntax oder Signale bspw. eines physikalischen Vorgangs gemeint. Aus diesen Daten werden Informationen durch Hinzugabe einer Interpretation, die diesen Daten jeweils einen Bedeutungsgehalt (Semantik) zuweist. Die jeweilige Bedeutung lässt sich nicht aus den Daten selbst ableiten, sondern ergibt sich aus der vom Menschen vorgenommenen Zuordnung einer Referenz oder einer physikalischen Größe. Beim Wissen wird die Information in einen Kontext gestellt, der eine zielgerichtete Informationsverarbeitung zulässt und auf dessen Basis letztlich eine Entscheidungsfindung zur Koordination von Handlungen möglich ist (Pragmatik).

Dieses Modell lässt sich auch auf die Wissenschaftsbewertung übertragen (Abb. 8.1), wie es in ähnlicher Weise unter dem Stichwort der Wissenschaftskommunikation bereits vorgestellt worden ist (Erlach 2022: 200). Datenbasis auf der untersten Ebene I sind die



Abbildung 8.1: Die drei Ebenen der Wissenspyramide zur Wissenschaftsbewertung, in Anlehnung an Erlach (2022: 200).

unterschiedlichen Einzel-*Ergebnisse* des wissenschaftlichen Arbeitens, die im Rahmen einer Datenerfassung gezählt oder gemessen werden können. Beispiele aus der Bibliometrie sind Zählgrößen wie die Anzahl der Publikationen von Autoren oder Institutionen sowie die Anzahl von Zitationen. Die Wissenschaftsindikatorik bietet aber noch weitere Zählgrößen wie die Anzahl der betreuten Dissertationsvorhaben oder die Anzahl von Forschungsprojekten eines bestimmten Typus an. Erfassbar sind aber auch dimensionsbehaftete Messgrößen wie eingeworbene Drittmittel in Euro. Formal betrachtet liegen mit diesen Zähl- und Messgrößen absolute Zahlen vor, auch in Form von Zeitreihen, die man auch als *Kennwerte* der Wissenschaftsbewertung bezeichnen kann.

Eine bewertende Bedeutung erhalten die Kennwerte auf der übergeordneten Informationsebene II durch ihre Verrechnung zu einer Information über die wissenschaftliche *Güte*. Dazu eignen sich zum einen dimensionsbehaftete oder auch dimensionslose relative *Kennzahlen* wie beispielsweise der *Journal-Impact-Factor* (JIF). Zum anderen hat die Bibliometrie auch kompliziert

strukturierte *Index-Werte* wie den *Hirsch-Faktor* (auch h-Index) zur Bewertung von Forscherleistungen entwickelt. Dieses datenbasierte Vorgehen scheint eine objektive Bewertung wissenschaftlicher Exzellenz zu ermöglichen, ohne dabei die einzelnen Ergebnisse inhaltlich zu berücksichtigen.

An der Spitze der Wissenspyramide steht die Wissensebene III, der als pragmatische Aufgabe die Wissensevaluierung durch Bewerten des wissenschaftlichen *Niveaus* zukommt. Dies kann durch einschlägige Fachexperten in Form von *Gutachten* im Peer-Review-Verfahren erfolgen. Dieses qualitative Vorgehen kann dabei keine objektivierenden Zahlenwerte, sondern bestenfalls intersubjektiv als richtig erachtete, qualitative Bewertungsaussagen liefern. Dafür erfolgt aber eine inhaltliche Betrachtung des zu evaluierenden wissenschaftlichen Wissens durch Prüfung der fachlichen Korrektheit (Verifikation der Begriffe, Modelle und Methoden) und im Fall angewandter Forschung auch eine Validierung des Nutzens. Fehlende Standards führen allerdings oft zu schlecht vergleichbaren und uneinheitlich strukturierten Gutachten, welche wiederum das Verfahren als sehr subjektiv und machtpolitisch missbrauchbar erscheinen lassen.

2. Über die Möglichkeit objektiver Wissenschaftsbewertung

Das Peer-Review-Verfahren ist immer wieder kritisiert worden, weil es mangels verbindlicher Bewertungskriterien im Prinzip nach Belieben subjektiv manipuliert werden kann. Insbesondere Nachwuchswissenschaftler beklagen daher, dass das Verfahren die Machtposition von Gatekeepern begünstigt. Daher fordern auch junge Wissenschaftsphilosophen in ihrem (natürlich gereviewten) Aufsatz *Is Peer Review a Good Idea?* die Abschaffung dieses klassischen Vorgehens der Qualitätssicherung (vgl. Heesen/Bright 2021). Umgekehrt möchte man es sich eventuell auch als etablierter Wissenschaftler in diesem System des wechselseitigen Begutachtens nicht mit den Kollegen durch zu harte

Kritik verscherzen – gerade weil sich qualitativ-inhaltliche Argumente auch in anonymen Verfahren zumindest vermutungsbasiert bestimmten Personen(kreisen) zuordnen lassen. Oder jemand lehnt eine stringent formulierte, sorgfältige Arbeit ab, weil er sich einem konkurrierenden Forschungsparadigma zurechnet und deshalb andere Ansätze mit deren spezifischen begrifflichen und methodischen Zugängen prinzipiell missbilligt. Wegen dieser Eigeninteressen stehen Gutachten unter dem Verdacht einer bloß subjektiven Bewertung wissenschaftlicher Exzellenz.

Die bibliometrische Bewertung wissenschaftlicher Exzellenz mit einem auf gezählten Forschungsergebnissen basierenden algorithmischen Bewertungsvorgehen bietet hier Abhilfe, weil es ein quantitativ-objektiviertes Gütemaß bereitzustellen verspricht. Allerdings gilt auch bei der bibliometrischen Leistungsmessung der alte Lehrsatz der Messtechnik »Wer misst, misst Mist«. Gemeint sind damit neben den im Alltag auftretenden, prinzipiell vermeidbaren Anwendungs- und Bedienungsfehlern die notwendigerweise auftretenden Messabweichungen (DIN 1319–1 1995: 12). Das beginnt erstens mit offensichtlichen, stochastischen Fehlern: In den Datenbanken tauchen Autoren in unterschiedlicher Schreibweise auf – dank der eindeutigen Autorenidentifikation mit der ORCID ein künftig vernachlässigbares Problem – oder manche Aufsätze sind unerwartet doch nicht automatisch erfasst worden – hier hilft dann nur manuelle Fleißarbeit der Autoren. Hinzu kommen zweitens bekannte systematische Fehler: Die Datenbanken berücksichtigen nicht alle Publikationen, sondern nur Aufsätze aus ausgewählten Zeitschriften, vorzugsweise englischsprachig. Damit sind zum einen Bücher – vor allem Monografien, aber auch Sammelbände – deutlich unterrepräsentiert bzw. tauchen als wissenschaftlich relevante Publikationsart gar nicht mehr auf. Gerade dies fördert auch den Trend zum kumulativen Promovieren, weil dadurch die Präsenz in den Datenbanken gesichert ist – und dies sogar mit mindestens drei Papern anstelle ‚nur' einer Monografie. Zum anderen übernehmen die privatwirtschaftlich geführten Datenbanken eine Gatekeeper-Funktion, indem sie den Zugang

mancher Zeitschriften zur Datenbank nach allenfalls intersubjektiv bewertenden Kriterien der Wissenschaftlichkeit je nach Interessenlage erschweren (können).

Drittens gibt es dann noch die in ihrem Auftreten und ihrem Ausmaß unbekannten systematischen Fehler: Dies „ist der Effekt, dass die Messung das zu messende Phänomen beeinflussen kann, die objektive Messung mithin nicht lange wirklich objektiv bleibt. Letztlich handelt es sich hier um ein Sozialexperiment mit Wissenschaftlern, das aufgrund der transparent gemachten Ergebnisfunktionen mit JIF und h-Index eine methodenbedingte *Reaktivität* aufweist. Es ist naheliegend, dass Forscher versuchen, ihre Bewertungsergebnisse gezielt – und unabhängig von etwaigen inhaltlichen Anforderungen – zu verbessern. Dabei tritt ein typisches Problem von Kennzahlensystemen auf: Wer das System durchschaut, kann es auch (in Grenzen) manipulieren. Mit Hilfe von Zitationstausch im engeren Kollegenkreis und Selbstzitaten von jeweils kennzahlrelevanten, dabei aber inhaltlich nur mäßig passenden Publikationen lässt sich das Messergebnis bezogen auf die eigentliche Messabsicht verfälschen“ (Erlach 2022: 200 f.). Insgesamt erweist sich das bibliometrische Bewertungsverfahren also als viel subjektiver, als es zunächst den Anschein hat.

Offenbar ist die simple Opposition subjektiv-manipulierendes Peer-Review-Verfahren versus objektiv-messendes Datenbank-Verfahren eher eine karikaturhafte Übertreibung, denn eine haltbare Entscheidungsgrundlage zur validen Auswahl eines Bewertungsverfahrens. Beide Vorgehensweisen sind bis zu einem gewissen Grade manipulierbar und interessengeleitet. Daher muss man sich bei beiden Verfahren um eine Minimierung negativer Auswirkungen auf die Bewertungsgüte Gedanken machen. Der Fokus auf eine Objektivierung der Bewertung mit Zahlen diskreditiert alle nach intersubjektiven Kriterien gestalteten, qualitativen Evaluierungen als bloß subjektiv, und als daher zu vermeiden. Es gibt jedoch keine harte Trennung zwischen objektiver und subjektiver Wissenschaftsbewertung, beides sind idealtypische Grenzfälle eines möglichen Verfahrensspektrums, das im Folgenden kurz umrissen werden soll.

3. Bewertungsverfahren im Spektrum von subjektiv bis objektiv

Der Grenzfall eines subjektiven Bewertungsverfahrens ist die *Beurteilung* in einem schriftlichen Arbeitszeugnis. Zahlreiche Verhaltensweisen und Handlungen, teilweise länger zurückliegend, werden nach abstrakten Kriterien summarisch ohne konkrete Verweise benotet und dann in relativ starre Standardformulierungen übertragen (Abb. 8.2, Zeile 1). Eine wissenschaftliche *Begutachtung* kann hingegen intersubjektiv nachvollziehbar sein, weil sie – gegliedert nach unterschiedlichen Aspekten wie Originalität der Ergebnisse, Angemessenheit der Methodik, Umfang des berücksichtigten Standes der Wissenschaft, formale Qualifikation des Autors/Antragstellers etc. – auf einen konkreten Text bezogene Urteile bietet. Diese münden dann in ein begründetes, eindeutiges Votum, ggf. mit Änderungsvorschlägen versehen (Abb. 8.2, Zeile 2). Daran lässt sich bereits erkennen, dass sich die Objektivität von Gutachten durch Vorgabe einer möglichst klar strukturierten und vollständigen Liste der Beurteilungsaspekte prinzipiell verbessern lässt.

Ein in der Industrie häufig genutztes Verfahren zur Bewertung komplexer (soziotechnischer) Systeme ist die *Nutzwertanalyse*

Nr.	Verfahren	Ermittlung & Beurteilung	Typ
1	Beurteilung	Zeugnis mit Noten nach abstrakten Eigenschaften & Standardformulierungen	subjektiv-qualitativ
2	Begutachtung	Gutachten mit Urteilen nach Aspekten & eindeutiges Votum, ggf. mit Änderungsvorschlägen	intersubjektiv-qualitativ
3	Nutzwertanalyse	Nutzwert-Punkte nach Bewertungskriterien & Addition nach Gewichtung der Punkte && Vergleich Gesamtnutzen mit Aufwand	intersubjektiv-qualitativ
4	Benotung	Teilnoten nach Aspekten *oder* Kriterien & Addition, ggf. nach Gewichtung	intersubjektiv-qualitativ
5	Benotung	Addition von Punkten für Teilergebnisse & Zuordnung zu Notenskala	intersubjektiv-quantitativ
6	technische Prüfung	Zählung / Messung nach Vorschrift mit Maßstab & Vergleich mit Sollwert	objektiv-quantitativ

Abbildung 8.2: Exemplarische Bewertungsverfahren in einer Abfolge von subjektiv bis objektiv.

(vgl. Rinza 1977). Dabei werden zunächst fallspezifisch qualitative Bewertungskriterien festgelegt und in der Regel durch paarweisen Vergleich in ihrer Relevanz relativ zueinander gewichtet (Abb. 8.2, Zeile 3). Für unterschiedliche Lösungsansätze einer Systemgestaltung wird anschließend durch Punktevergabe für jedes einzelne Bewertungskriterium je nach Erfüllungsgrad (häufig auf einer Skala von 1 bis 5) ein Nutzwert ermittelt. Der Gesamt-Nutzwert einer Lösung ergibt sich dann als gewichtete Summe der Einzelwerte. Gewichtung und Bewertung erfolgen gemeinsam in einer Gruppe aus Fachexperten und Entscheidern, ist also intersubjektiv angelegt. Durch die schrittweise Bewertung nach Kriterien wird eine subjektiv-intuitive Bewertung ‚aus dem Bauch heraus' vermieden. Die Erfahrung zeigt, dass allein durch die systematische Diskussion der Gründe für eine Bewertung subjektive Fehleinschätzungen oder sachfremde Bewertungsgründe reduziert werden (Erlach 2020: 321). Eine Besonderheit des Verfahrens ist die Möglichkeit, die Nutzwerte jeder Lösung den jeweiligen zur Umsetzung erforderlichen Aufwänden in einem Diagramm gegenüberzustellen, wobei der Nutzwert auf der einen Achse und der Aufwand auf der anderen aufgetragen werden. Die relativ beste Vorzugs-Lösung in diesem *zweidimensionalen* Bewertungsverfahren ist dann in der Regel weder die mit dem höchsten Nutzwert – der notorisch mit den höchsten Aufwänden verbunden ist – noch die mit den niedrigsten bzw. gar keinen Aufwänden, da in aller Regel mit vergleichsweise aufwandsarmen Maßnahmen bereits ein nennenswerter Nutzen erreichbar ist. Die Vorzugs-Lösung kombiniert dann jene Aufwände, die viel Nutzen versprechen und ignoriert jene hohen Aufwände, die nach der Regel des abnehmenden Grenznutzens, zwar zu einer perfekten Lösung führen, den Aufwand aber als unangemessen erscheinen lassen.

Bei der *Benotung* in der Lehre (Schule und Universität) kann man zwei grundsätzlich unterschiedliche Vorgehensweisen unterscheiden – die qualitative Vergabe von Teilnoten und die quantitative Punktbewertung von Teilergebnissen. Die Benotung beispielsweise einer philosophischen Hausarbeit nach unterschiedlichen Aspekten hat dabei gewisse Ähnlichkeiten zu einem

Gutachten, wenn auch nicht so elaboriert und oft auch nicht so systematisch. Eine stärker formalisierte Benotung nach gewichteten Kriterien weist Ähnlichkeiten zur Nutzwertanalyse auf (Abb. 8.2, Zeile 4). Ein solches Vorgehen wird derzeit von mehreren Professoren für von Mitarbeitern des Fraunhofer IPA erstellten Dissertationen in den Ingenieurswissenschaften genutzt. Es weist 16 Bewertungskriterien in den vier Kategorien Anwendungsbezug, wissenschaftliche Vorgehensweise, formale Güte und persönlicher Beitrag auf (Erlach 2025: 55).

Die Benotung von Prüfungsleistungen, bei denen das Erreichen von Teilergebnissen, wie beispielsweise einer mathematischen Berechnung oder der richtigen Auswahl in einem Multiple-Choice-Test, eindeutig festgestellt werden kann, beinhaltet als subjektive Elemente lediglich die relative Gewichtung der Teilergebnisse über die Punktezuordnung zueinander (Abb. 8.2, Zeile 5). Dieses Verfahren setzt einen stark formalisierten Prüfungsgegenstand voraus, der in der Mathematik meist angemessen ist, in anderen Fachgebieten aber erst durch Abstraktion bzw. Vereinfachung hergestellt werden muss. Die qualitative Güte von Lösungen lässt sich so nicht bewerten.

Der *technischen Prüfung* liegt eine objektiv-quantitative Messung zu Grunde, bei der festzustellen ist, „inwieweit ein Prüfobjekt eine Forderung erfüllt" (DIN 1319–1 1995: 6). Messgröße ist eine physikalische Größe, die durch Anwendung eines physikalischen Messprinzips mit einer Messmethode oder auch durch einfaches Zählen (Zählgröße) ermittelt wird. Dabei wird der Maßstab einer physikalischen Einheit verwendet, der jedoch rein konventioneller Natur ist und das Messergebnis als solches nicht beeinflusst. Entscheidend ist, dass es zur Prüfung des Vergleichs einer ‚Forderung' bedarf, also ein zu erreichender Sollwert vorgegeben oder vereinbart sein muss (Abb. 8.2, Zeile 6). Hier kommt allerdings eine subjektive Komponente in die Prüfung, da der Sollwert keine bloße Konvention ist, sondern eine wählbare Anforderung an die Qualität darstellt.

Für die Anwendung der Bibliometrie in der Wissenschaftsbewertung heißt das, dass es nicht ausreichend ist, nur die

Kennzahlen (Faktoren und Indizes) gewissermaßen unsortiert zu ermitteln. Daher ist es auch üblich, sie nach Größe zu sortieren und dann dem höchsten Wert die jeweils höchste wissenschaftliche Qualität zuzusprechen. Letztlich ist dies aber eine implizite, scheinbar plausible und daher nicht weiter zu rechtfertigende Sollwertfestlegung, die sich um eine explizite Vorgabe angemessener Ziele drückt. Besonders deutlich wird diese Leerstelle beim Vergleich zweier Wissenschaftler mit unterschiedlicher Dauer ihrer Laufbahn. Wenn der eine am Ende und der andere am Anfang seiner Karriere steht, dann sind sie bei einer identischen bibliometrischen Datenlage ganz offensichtlich nicht gleich zu bewerten (vgl. dazu den Beitrag „Perils and Promises" von H. Andersen in diesem Band).

Die Vorgabe von Sollwerten hätte zudem den angenehmen Nebeneffekt, dass auch quantitative Auswüchse nach oben, wie sie nur durch eine große Mitarbeiteranzahl, nicht aber allein durch eigene Leistung erreichbar sind, tendenziell weniger positiv bewertet werden würden. Voraussetzung wäre, wie das Beispiel der Nutzwertanalyse zeigt, ein breiter disziplinärer Diskurs über eben diese Sollwerte, der durch eben diese Reflexionsleistung allein auch ohne breiten Konsens das Bewertungsverfahren selbst oder zumindest das Verständnis seiner Gültigkeitsbedingungen verbessern würde. Auch wenn dieser Austausch in der notwendigen Breite utopisch ist, könnten sich sicherlich Wissenschaftsinstitutionen für bestimmte Anwendungsbereiche, z.B. bei bestimmten Förderformaten, über Sollwerte Gedanken machen, um das bibliometrische Verfahren nicht ganz zielblind laufen zu lassen. Denkbar wäre es, beim Hirsch-Faktor lediglich zu prüfen, ob er über oder unter einem Grenzwert liegt.

4. Methoden der Qualitätssicherung in der Industrie

Eine wie auch immer geartete Wissenschaftsbewertung setzt ein entsprechendes Verständnis der *Qualität* von Wissenschaft voraus. In der Wissenspyramide lassen sich die beiden Ansätze des

quantitativen Berechnens der wissenschaftlichen *Güte* einerseits sowie das qualitative Bewerten des wissenschaftlichen *Niveaus* andererseits unterscheiden (Abb. 8.1). Der Ausflug in die Messtechnik hat gezeigt, dass Qualität zudem einen Vergleich mit einer Forderung, einem Sollwert, voraussetzt. In der Industrie ist dafür das Qualitätswesen zuständig. Dabei gilt: „Qualität beschreibt die Übereinstimmung (*Konformität*) eines Produktes, eines Prozesses oder einer Tätigkeit mit vorgegebenen Forderungen" (Brüggemann et al 2024: 4). Interessanterweise hat sich das Qualitätswesen in der Industrie im Laufe des letzten Jahrhunderts langsam von einer rein quantitativen zu einer stark qualitativ geprägten Vorgehensweise weiterentwickelt – also gerade umgekehrt zur Wissenschaftsbewertung, die die rein qualitative Begutachtung von Inhalten erst später durch das quantitative Zählen von Ergebnissen ergänzt hat.

Bei der Entwicklung des Qualitätswesens kann man, stark vereinfachend, vier Stufen unterscheiden (Abb. 8.3; Brüggemann et al 2024: 5 ff.; Jakoby 2022: 27 ff.): Im Unterschied zum Handwerk ist erst durch die Massenfertigung am Beginn der Industrialisierung eine Wiederholgenauigkeit bei der Herstellung von Teilen und Produkten notwendig geworden. Einerseits sollen die Wiederholteile ja austauschbar sein und andererseits sollen im arbeitsteilig zergliederten Herstellungsprozess zueinander passende Teile produziert werden. Aufgabe der *Qualitätssicherung* (Stufe 1) ist es, schlechte Teile (den Ausschuss) durch kontrollierende, unabhängige Qualitätsprüfabteilungen auszusortieren. Dies entspricht dem Funktionsmeisterprinzip, wie es Frederick W. Taylor vorgeschlagen und schließlich 1911 publiziert hat. Grundannahme der Qualitätssicherung ist, dass einzelne Menschen durch Unachtsamkeit Fehler erzeugen, die dann nachträglich auszusortieren sind. Verbesserungen sind bei dieser Sichtweise eigentlich nur durch intensivere Überwachung der Mitarbeiter und ihrer Arbeitsweise zu erreichen. Qualität wird (scheinbar) rein individuell erzeugt.

Relativ schnell hat man gemerkt, dass dieses Vorgehen, im Nachhinein wie ein Beobachter passiv den Ausschuss zu entsorgen, sehr teuer ist. Da Material und alle bisherigen Arbeitsaufwände

Methode	charakteristische Merkmale	Vertreter
Qualitätsprüfung (sortierend)	Aussortieren schlechter Teile durch kontrollierende Qualitätsprüfabteilungen *Fokus: Einzelne **Menschen** erzeugen Fehler*	ab 1910 Frederick W. Taylor
Qualitätslenkung (steuernd)	Stichprobenkontrolle mit statistischer Prozesslenkung: „Kontroll-Karten“ *Fokus: Fehler haben **systematischen** Charakter*	ab 1930 Walter A. Shewhart
Qualitätsmanagement (vorbeugend)	Kontinuierliche Verbesserung der fehlererzeugenden Produktionsprozesse: Null-Fehlerstrategie *Fokus: **Prozesse** provozieren Fehler*	ab 1940 W.E. Deming Ph.B. Crosby
Qualitätsmanagement-system (integrierend)	Einbezug aller Unternehmensprozesse (Vertrieb, Beschaffung, Produktentwicklung, etc.): Total-Quality-Management (TQM), Auditierung *Fokus: Qualität als **Unternehmensziel***	ab 1950 A.V. Feigenbaum

Abbildung 8.3: Historische Entwicklung der Methoden des Qualitätswesens in der Industrie.

verloren gehen, gilt der Ausschuss in der ursprünglich von Toyota entwickelten, mittlerweile weltweit verbreiteten Produktionsmethodik des ‚Lean Production' als die schlimmste aller Verschwendungsarten (Erlach 2020: 128). Aufgabe der seit den 1930iger Jahren entwickelten *Qualitätslenkung* (Stufe 2) ist es, zunächst mit Hilfe der von Walter A. Shewhart eingeführten ‚Kontroll-Karten' die entdeckten Fehler mit statistischen Methoden zu untersuchen und zu bewerten. Die dazu erforderlichen Stichprobenkontrollen senken zum einen den Prüfaufwand gegenüber der 100%-Kontrolle der einfachen Qualitätssicherung und ermöglichen zum anderen eine Fehlerreduktion durch die statistische Analyse des Produktionsprozesses. Grundannahme dieser statistischen Prozesslenkung ist, dass Fehler einen systematischen Charakter haben und daher nach ihrer Identifikation durch Prozessanpassung reduziert werden können. Der Fokus verschiebt sich vom Produkt auf die Produktion und dort auf die Befähigung von Technik und Mitarbeitern durch Qualifikation.

Es ist das Verdienst des Amerikaners W.E. Deming, den Gedanken, dass Qualität nicht geprüft, sondern erzeugt werden muss, in Richtung *Qualitätsverbesserung* systematisch weiterentwickelt zu haben. Da das diesbezügliche Interesse der amerikanischen Industrie in den 1940iger Jahren nicht sehr hoch war, hat er

zunächst den Aufstieg der japanischen Industrie mit seiner Idee des ‚Kontinuierlichen Verbesserungsprozesses (KVP)' geprägt. Heute noch bekannt und vielfach angewendet ist der sogenannte PDCA-Zyklus, wobei das Akronym für die vier Phasen ‚Plan-Do-Check-Act' der Qualitätsplanung und -verbesserung steht.

Erst zwanzig Jahre später hat Ph.B. Crosby mit der ‚Null-Fehlerstrategie' in der amerikanischen Industrie Aufmerksamkeit erzeugt. Hierbei geht es darum, das Machen von Fehlern durch geeignete Prozessgestaltung zu erschweren oder auszuschließen. Die europäische Industrie hat sich noch deutlich später mit diesen Ansätzen auseinandergesetzt – letztlich aufgeschreckt durch den ‚überraschenden' (und bedrohlichen) Erfolg Japans.

Aufgabe des *Qualitätsmanagements* (Stufe 3) ist es, mit geeigneten Management-Methoden systematisch eine kontinuierliche Verbesserung der fehlererzeugenden Produktionsprozesse zu erreichen. Ein anschauliches Beispiel für eine der Methoden ist das fischgrätenartige Ursache-Wirkung-Diagramm (nach dem Erfinder auch Ishikawa-Diagramm genannt), bei dem die sechs Äste (‚Gräte') möglicher Ursachenbereiche Mensch-Maschine-Milieu-Material-Methode-Messung (Alliteration als Gedächtnisstütze) in einen Hauptpfeil (‚Wirbelsäule') zum problematischen Fehler führen. Grundannahme des Qualitätsmanagements ist, dass (schlechte) Prozesse Fehler provozieren und daher verbessert werden müssen, um letztlich eine höhere Produktqualität zu erreichen. In dieser Sichtweise ist der Einbezug der Mitarbeiter in die Prozessgestaltung eine Voraussetzung für den Erfolg und Qualität ein Ergebnis der Teamarbeit von Ingenieuren und Werkern.

Man könnte hier einen Endpunkt der Entwicklung vermuten, aber Überlegungen von A.V. Feigenbaum bereits Mitte der 1950iger Jahre zum ‚Total Quality Control' haben die Grundlage für den nächsten Schritt gelegt, der vor allem in den 1980igern mit der Normenfamilie ISO 9000 ff. begonnen hat. Aufgabe des in diesen Normen beschriebenen (und beständig weiterentwickelten) *Qualitätsmanagementsystems* (Stufe 4) ist es, alle Unternehmensprozesse vom Vertrieb über Produktentwicklung bis

hin zur Beschaffung in die systematische Qualitätsverbesserung einzubeziehen. „Ein *Qualitätsmanagementsystem (QMS)* ist eine Einheit zusammenwirkender Methoden und Dokumente zur Planung und Steuerung der Prozesse in einer Organisationseinheit, die die Erfüllung der Produktanforderungen sicherstellen" (Jakoby 2022: 156). Grundannahme der Einführung von Total-Quality-Management-Systemen (TQM-Systeme) ist, dass Qualität als eigenständiges Unternehmensziel angesetzt wird und der ökonomische Gewinn nur als eine der angestrebten Ergebnisgrößen auftaucht. Ein sehr umfassendes Unternehmensmodell dieser Art ist das EFQM-Modell der ‚European Foundation for Quality Management'. In dieser Sichtweise sind die fünf sogenannten ‚Befähiger' – Führung, Strategie, Mitarbeiter, Partner und Ressourcen, Prozesse und Produkte – Voraussetzung zur Erfüllung der vier zentralen Ergebnisdimensionen eines Unternehmens, nämlich Kunde, Mitarbeiter, Gesellschaft und viertens die klassischen Schlüsselergebnisse (Key Performance Indicators – KPI), die als einzige Ergebniskategorie quantitativ mess- und zählbar sind.

Qualität ist also in TQM-Systemen nur noch zu einem kleinen Anteil direkt quantitativ messbar. In den Fokus gerückt ist hingegen die *Auditierung*, bei der externe oder auch interne Experten eine fragebogen- und checklistengeleitete (Selbst-)Bewertung des Unternehmens in den jeweils definierten Dimensionen durchführen. Als Skala dienen häufig Reifegradpunkte, und der jeweilige Zielwert resultiert aus Vorgaben der Unternehmensleitung, meist angelehnt an branchentypische Referenzwerte. Man könnte also sagen, dass hier ein systematisch strukturiertes Gutachten erstellt wird.

In gewisser Weise vollzieht hier die Wirtschaft eine Gegenbewegung zur Wissenschaft: weg vom bloßen Zählen und Messen (Prüfung) und hin zum mehrdimensionalen qualitativen Bewerten durch Experten (Auditierung). Möchte man aus der Entwicklung des Qualitätswesens in der Industrie etwas lernen, dann kann man sagen, dass es vielleicht gar nicht um eine möglichst objektive Wissenschaftsbewertung gehen sollte, sondern vielmehr um

eine Verbesserung der Wissenschaft. Dafür müssten dann Methoden des Qualitätsmanagements für die Wissenschaft entwickelt werden im Sinne eines Wissenschaftsmanagements, dem es nicht nur um die Verwaltung von Forschungsgeldern geht.

5. Qualitätsmanagement für die Wissenschaft?

Bereits in den 1960igern hat Solla Price bei seiner statistischen Betrachtung des Wissenschaftssystems als Ganzem ermittelt, dass sich die Zahl der Wissenschaftler seit 1660 – beginnend mit der Gründung der ersten wissenschaftlichen Fachgesellschaften und der ersten wissenschaftlichen Zeitschriften – ungefähr konstant in einem Zeitraum von zehn Jahren verdoppelt (Solla Price 1974: 17). Schreibt man dies fort, entspräche in mittlerweile 150 Jahren die Zahl der Wissenschaftler der doppelten Bevölkerungszahl (ebd.: 30) – oder anders gesagt: es muss inzwischen zum deutlichen Abbremsen des Wachstums gekommen sein.

Matthias Kölbel hat in einer Untersuchung für Deutschland einen zunächst hyperbolischen anstatt exponentiellen Verlauf des Wachstums von Professorenstellen sowie der Anzahl an Universitäten gefunden, was daran liegen mag, dass die Wissenschaft durch die ‚Verwissenschaftlichung der Gesellschaft' ihre Wachstumsbedingungen zusätzlich zum exponentiellen Wachstum gleichzeitig noch steigert. Der Umbruch in die Stagnation zu Beginn der 1980iger Jahre ist dadurch nur umso stärker. Gleichzeitig stagnieren die anteilig am BIP gemessenen öffentlichen Forschungsmittel seit den 1970iger Jahren. Die Aufwendungen für Forschung und Entwicklung (FuE) der Industrie hingegen brechen zeitverschoben in den 1990igern deutlich ein (Kölbel 2002: 115 f.).

Bei einer Stagnation der Anzahl an Wissenschaftlern und der Forschungsmittel lässt sich qualitatives Wachstum der Wissenschaft kaum noch durch bloße Erhöhung der Ergebnis-Quantität erreichen – also dadurch, dass bei deutlich mehr Publikationen auch automatisch etwas mehr hochwertige

Publikationen entstehen. Dabei gilt wohl der statistische Zusammenhang, dass die Gesamtzahl der Publikationen mit dem Quadrat der Zahl relevanter Publikationen wächst (Solla Price 1974: 67). Obwohl Solla Price davon ausgeht, dass man auch bei einer von ihm vermuteten Korrelation quantitativer und qualitativer Gewichtung eines Forschers dessen individuellen Grad der Bedeutung nicht wirklich objektiv messen kann (ebd.: 88), war er doch einer der Ideengeber für die bibliometrische Vermessung (und Bewertung) der Wissenschaften.

Es entbehrt nicht einer gewissen Ironie, dass aus der von ihm zugleich betonten, zwangsläufigen Begrenzung des quantitativen Wachstums (relativ zur Bevölkerungszahl) früher oder später nennenswerte qualitative Fortschritte gar nicht mehr über die schiere Masse an Publikationen erreicht werden können. Die maximale Anzahl an Publikationen für ein Forscherleben sieht er bereits in den 1960igern bei einer Zahl von 1.000 fixiert; und es erscheint plausibel, dass sich das ohne weitere Co-Autoren (also noch mehr Forscher) nicht weiter steigern lässt. Vielleicht lässt sich durch die Nutzung von LLMs wie ChatGPT das je Autor produzierte Textvolumen nochmals deutlich steigern, aber es ist doch fraglich, ob das ohne Verluste der Durchschnitts-Qualität im Sinne eines Neuigkeitsgehaltes überhaupt möglich ist.

Ein qualitätsorientiertes Wissenschaftsmanagement sollte sich also um die qualitative Verbesserung der Wissenschaft und weniger um das bloße Selektieren von besonders guten Individuen nach Zahlen bemühen. Die wissenschaftliche Leistung wird dann nicht ausschließlich individuell bewertet, sondern auch prozessual mit dem jeweiligen institutionellen Kontext des einzelnen Wissenschaftlers im Team. Vermutlich hat vor gut 20 Jahren der Bologna-Prozess dazu geführt, dass zahlreiche Universitäten ein Qualitätsmanagementsystem eingeführt haben.

Gezeigt hat sich das in der Regel in der Evaluation der Lehre, insbesondere im Zusammenhang mit der Akkreditierung von Studiengängen, wie eine kleine Internetrecherche auf den Websites der Universitätsverwaltungen unter dem Stichwort ‚Qualitätsmanagement' schnell zeigt. Ein sehr schön aufbereitetes Beispiel ist

das QM-System heiQUALITY der Universität Heidelberg. Das System dient der „Sicherung und Weiterentwicklung eines hohen Qualitätsbewusstseins […] gelebt in einer entsprechenden Qualitätskultur", wobei „der Begriff des Qualitätsmanagements primär im Sinne einer kontinuierlichen Qualitätssicherung und -entwicklung zu verstehen" ist (Uni Heidelberg 01). Dementsprechend ist auch der PDCA-Zyklus dargestellt. Mit der „Etablierung und Festigung einer lebendigen Qualitätskultur in der Betreuung der Doktoranden" sollen fakultätsübergreifend Qualitätsstandards im Promotionswesen gehoben werden (Uni Heidelberg 02). Dadurch nähert sich das Qualitätsmanagement aus der Lehre kommend der Forschung an. Die Forschung selbst ist allenfalls proklamatorisch, kaum aber durch konkrete Angaben von den Qualitätsmanagementansätzen ‚betroffen'.

Deutlich über diese der Lehr-Evaluation entsprungenen QM-Systeme hinaus geht ein etwas verstecktes (nicht unter Qualität eingeordnetes) Dokument der Universität Utrecht aus dem akademischen Jahr 2023/24. Unter dem Titel *Vision on Recognition and Rewards* wird das ‚Utrecht TRIPLE model' vorgestellt (einen früheren Entwurf gibt es von 2021) und mit einer fünfblättrigen Lotusblüte, deren Zentrum der wissenschaftliche ‚Impact' ist, visualisiert. Grundlage für die drei Arbeitsfelder ‚Education', ‚Research' und ‚Professional performance/organisation' sind die beiden Basisblätter ‚Team' und ‚Leadership', womit der Fokus sehr deutlich weg vom Messen hin zur Zusammenarbeit, die zu Anerkennung und Belohnung führen soll, gesetzt ist.

Eines der Prinzipien für diesen Veränderungsprozess (‚movement') lautet ‚Quality over quantity', denn wie schon im industriellen Bereich ist die Qualität das übergeordnete Ziel: „The focus is on what makes our work valuable". Die neue Sichtweise wird im Folgenden sehr gut zum Ausdruck gebracht: „This means recognition and rewards is not just about replacing one set of assessment criteria with a new checklist; it is about a new way of looking at things. Instead of only looking at *what* someone did, we will also look at *how* they did it (team spirit, leadership), and based on *what objective* (impact). Qualitative indicators, such as substantiated

narratives, are very suited to this. Quantitative indicators can support the narrative. However, these indicators can never be used in isolation and they are means to an end, not ends in themselves" (Utrecht University 2023: 3). Bibliometrische Daten sind also nur noch zu verwenden, wenn damit ein konkret benanntes, gemeinsames Ziel adressiert worden ist. Analog zum TQM ist damit ein ‚Team Science Impact'-System konzeptioniert und eingeführt. Ob sich dadurch die (lokale) akademische Kultur ändern kann, muss sich jedoch erst noch zeigen.

6. Qualitätsbewertung oder Qualitätsmessung in der Philosophie

Was heißt das nun für die Philosophie? Im Gegensatz zu den Experimentalwissenschaften wird ja hier nicht in Teams gearbeitet, jedenfalls ist es weder notwendig, wie bei den Großforschungseinrichtungen, noch auch üblich. Dem Klischee entspricht eher der Gelehrte in seiner Studierstube, der Bücher liest, Zettelkästen füllt und Texte – meist Bücher – schreibt. Trotzdem ist ein Wandel erkennbar. Zeitschriftenartikel werden relativ zu den Büchern immer wichtiger und die Drittmittelforschung begünstigt Co-Autorenschaften, vor allem bei interdisziplinären Projekten, zumeist im Bereich der praktischen Philosophie. Wenn Solla Price recht hat, und sich Wissenschaft unabhängig von ihren Inhalten quantitativ weiterentwickelt, dann würden Mehrautorenschaften auch in der Philosophie zum Normalfall werden und eine teamorientierte Qualitätsbewertung geeignet erscheinen.

Bei genauerer Betrachtung kennt die Philosophie die Teamarbeit eigentlich schon seit ihrem Anbeginn, und damit länger als die moderne Wissenschaft, – nämlich mit dem Phänomen der Schulen-Bildung. Das mag nicht immer geglückt sein – es ist jetzt hier nicht der Ort, das philosophiehistorisch näher zu untersuchen und zu bewerten. Aber gerade das zeigt ja auch, dass qualitativ gelungenes Teamwork mit integrativer (Selbst-)Führung innerhalb einer teilweise informellen, teilweise institutionell

verbundenen Gruppe von Philosophen zwar nicht selbstverständlich, aber doch auch erstrebenswert ist.

Der zunehmende Fokus auf Qualitätsmessung mit bibliometrischen Verfahren auch in der Philosophie hat – wie in anderen Wissenschaften – zu paradoxen Effekten geführt. Wendet man, wie in der Bibliometrie üblich, ein quantitatives Verfahren ohne Maßstab an, dann ist es nicht nur qualitätsblind, sondern sogar tendenziell qualitätsschädigend, weil die erforderliche Maximierung des Outputs Wissenschaftler zur Reduktion der Qualität wenn nicht zwingt, so doch animiert. Die daher zu Recht geforderte Rückbesinnung auf qualitative Bewertungsverfahren sollte jedoch nicht ausschließlich auf klassische Review-Verfahren setzen, die wegen ihres hohen Aufwandes und ihrer fehlenden Standardisierung (vgl. dazu den Beitrag „Epistemic criteria for peer review in philosophy of science" von N. Hangel in diesem Band) ebenfalls deutliche und häufig kritisierte Mängel aufweisen.

Während der durch bibliometrische Qualitätsmessung erzeugte qualitätserodierende Effekt in der Physik durch den hohen mathematischen Formalisierungsgrad oder auch den hohen apparativen Aufwand ein gewisses Gegengewicht bekommt, ist ein solches beispielsweise in Medizin, Psychologie und Lebenswissenschaften trotz des konkreten Gegenstandsbezugs offenbar gar nicht vorhanden, wie die Replikationskrise zeigt. In der Philosophie fehlt das Gegengewicht empirischer Gegebenheiten schon aus grundsätzlichen Gründen, weil sie es weniger mit Fakten und mehr mit Argumenten zu tun hat. Ob eine Argumentation als gelungen gilt oder nicht, ist der eigentliche philosophische Gehalt und nicht eine Frage disziplinübergreifender Qualitätsbewertung. Dadurch erscheint eine Qualitätsmessung in der Philosophie noch problematischer.

7. Resümee

Dieser Beitrag sollte Ansatzpunkte zur Qualitätsbewertung in der Wissenschaft in zwei grundsätzlichen Hinsichten aufzeigen:

zunächst auf der elementaren Ebene der Bewertungsverfahren und sodann auf der systemischen Ebene der Qualitätssicherung.

Die Betrachtung von sechs grundlegenden Bewertungsverfahren zeigt, dass quantitative Verfahren für ihre ‚objektive' Messung durch einen Maßstab mit normativer Sollwert-Vorgabe gewissermaßen eingehegt werden können, damit sie nicht zu blindlings laufenden Maximierungen führen. Qualitative Verfahren können umgekehrt intersubjektiv validiert werden, indem sie mit Kriterien in unterschiedlichem Grad formalisiert werden.

Ein Blick auf die historische Entwicklung der Methoden des Qualitätswesens in der Industrie zeigt, dass hier eine zum Wissenschaftsbetrieb gegenläufige Entwicklung von rein quantitativen Qualitätsprüfungsmethoden hin zu qualitativ arbeitenden Qualitätsmanagementsystemen stattgefunden hat.

Dieser Trend ist keineswegs selbstverständlich, da man der gewinnorientierten Industrie mit Recht eine vom Controlling getriebene, rein zahlenbasierte Leistungsbewertung mit KPIs unterstellt – und dies auch unproblematisch erachtet im Unterschied zu einer ‚wahrheitsgetriebenen' Wissenschaft, deren primärer Bewertungsmaßstab nicht die Höhe der Forschungsgelder ist (? – oder sein sollte). Aber ein Jahrhundert Produktionspraxis und Produktionsforschung haben gezeigt, dass Qualität eine sensible Größe ist, die sich durch rein quantitative Methoden nicht erreichen lässt. Für die Wissenschaft wird das kaum anders sein, die Eule der Minerva ist zahlenblind.

Literatur

Andersen, Hanne (2026) „Perils and Promises: How to Regain Control of Metrics in Academia", in: Mößner, N.; Erlach, K. (Hg.), Vermessene Philosophie – Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum. Ubiquity Press, S. 295–317.

Aamodt, Agnar; Nygård, Mads (1995) „Different roles and mutual dependencies of data, information, and knowledge – An AI perspective on their integration", Data & Knowledge Engineering 3, S. 191–222. https://doi.org/10.1016/0169-023X(95)00017-M.

Brüggemann, Holger; Bremer, Peik; Zischka, Stefan (2024) Grundlagen Qualitätsmanagement. Von den Werkzeugen über Methoden zum TQM. Vieweg, https://doi.org/10.1007/978-3-658-43563-9.

DIN 1319-1 (1995) Grundlagen der Meßtechnik. Teil 1: Grundbegriffe. Beuth, https://dx.doi.org/10.31030/2713411.

Eliot, Thomas Stearns (1934) The Rock. Faber & Faber, https://archive.org/details/in.ernet.dli.2015.3608.

Erlach, Klaus (2020) Wertstromdesign. Der Weg zur schlanken Fabrik. Springer, https://doi.org/10.1007/978-3-662-58907-6.

Erlach, Klaus (2022) „Die Industrialisierung des Wissenschaftsbetriebs. Effiziente Erzeugung, Bereitstellung und Bewertung von Wissen", in: Mößner, N.; Erlach, K. (Hg.), Kalibrierung der Wissenschaft. transcript, S. 183–206. https://doi.org/10.14361/9783839462102-008.

Erlach, Klaus (2025) Propädeutik der Technikwissenschaften. Universität Stuttgart. https://doi.org/10.18419/opus-16595.

Hangel, Nora (2026) „Epistemic criteria for peer review in philosophy of science: a qualitative inquiry into a social epistemic practice", in: Mößner, N.; Erlach, K. (Hg.), Vermessene Philosophie – Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum. Ubiquity Press, S. 247–283.

Heesen, Remco; Bright, Liam Kofi (2021) „Is Peer Review a Good Idea?", in: The British Journal for the Philosophy of Science, S. 635–663. http://doi.org/10.1093/bjps/axz029.

Holzer, Angela (2022) „Die Vermessbarkeit der Wissenschaft. Digitalisierung, wissenschaftliches Publizieren, Verhaltenstracking und Wissenschaftsbewertung", in: Mößner, N.; Erlach, K. (Hg.), Kalibrierung der Wissenschaft. transcript, S. 163–182. https://doi.org/10.14361/9783839462102-007.

Jakoby, Walter (2022) Qualitätsmanagement für Ingenieure. Eine Einführung in die qualitätsorientierte Gestaltung von Produkten, Prozessen und Organisationen. Vieweg, https://doi.org/10.1007/978-3-658-36677-3.

Kölbel, Matthias (2002) „Das Wachstum der Wissenschaft in Deutschland 1650-2000", in: Parthey, H.; Spur, G. (Hg.): Wissenschaft und Innovation. Wissenschaftsforschung Jahrbuch 2001. Gesellschaft für Wissenschaftsforschung, S. 113–128. www.wissenschaftsforschung.de.

Lucky, R. (1989) Silicon dreams. Information, man, and machine. St. Martin's Press, https://lccn.loc.gov/89030091.

Mittermaier, Bernhard (2026) „Peer Review, Bibliometrie, Altmetrics – gibt es einen Königsweg bei Evaluierungen“, in: Mößner, N.; Erlach, K. (Hg.), Vermessene Philosophie – Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum. Ubiquity Press, S. 139–167.

Retzlaff, Eric (2022) „Wer bewertet mit welchen Interessen wissenschaftliche Publikationen? Eine Skizzierung des Einflusses kommerzieller Interessen auf die Forschungsoutput-Bewertung“, in: Mößner, N.; Erlach, K. (Hg.), Kalibrierung der Wissenschaft. transcript, S. 139–162. https://doi.org/10.14361/9783839462102-006.

Rinza, Peter; Schmitz, Heiner (1977) Nutzwert-Kosten-Analyse. Eine Entscheidungshilfe zur Auswahl von Alternativen unter besonderer Berücksichtigung nicht monetärer Bewertungskriterien. VDI-Verlag, https://doi.org/10.1007/978-3-642-51504-0.

Solla Price, Derek J. de (1974 [engl. 1963]) Little Science, Big Science. Von der Studierstube zur Großforschung. Suhrkamp, https://doi.org/10.7312/pric91844.

Uni Heidelberg 01. https://www.uni-heidelberg.de/de/universitaet/qualitaetssicherung-entwicklung/das-qm-system-heiquality, vom 11.12.2024.

Uni Heidelberg 02. https://www.uni-heidelberg.de/de/universitaet/qualitaetssicherung-entwicklung/das-projekt-heidocs, vom 11.12.2024.

Utrecht University: UU Vision Recognition and Rewards_2023.pdf, https://www.uu.nl/en/research/open-science/tracks/recognition-and-rewards, vom 11.12.2024.

CHAPTER 9

Epistemic Criteria for Peer Review in Philosophy of Science: A Qualitative Inquiry into a Social Epistemic Practice

Nora Hangel

1. Introduction

We can understand peer review as a special case of interpersonal social epistemology. What is different to other interpersonal interactions of critical scrutiny is that the criticism is blinded – double blinded in most contexts of philosophy – and organized in a special setting, since one to three referees assess a manuscript of one author.[1] Moreover, peer review is part of a social-epistemic institutionalized process of criticism (see Longino 1990; 2002). While this article investigates peer review in the first sense, its theoretical framing is shaped by the second, i.e., the institutionalized role of peer review by inquiring into the question of the role of referees for self-correction and innovation in philosophy.

How to cite this book chapter:
Hangel, N. 2026. Epistemic Criteria for Peer Review in Philosophy of Science: A Qualitative Inquiry into a Social Epistemic Practice. In: Mößner, N. and Erlach, K. (eds.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* Pp. 247–283. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.k. License: CC BY 4.0

For interpersonal social epistemology, the critical dimensions that impact the relationship between editors, referees, and authors include not only epistemic criteria but also scientific integrity, possible biases, and the positionality of the reviewers. The core task in the interpersonal review process is to assess arguments, rigor, and clarity while paying attention to how quality assessment works for different epistemic styles.[2] Epistemic quality control is thus related to the epistemic criteria applied within the situated context of the theoretical and methodical framing of a manuscript. However, one of the epistemic criteria referees look for is originality or novelty, which can activate individual biases of more conservatively minded or cautious referees.

The editor's role is to balance pragmatic and epistemic interests in the process of interpersonal peer review but editors also shape questions concerning new directions in a discipline through their impact on journals. This aspect of the process involves deliberation between the pragmatic aim of publication and organized criticism within a discipline (see Jukola 2017; Longino 1990). Depending of the decisions made in both interpersonal and institutionalized processes of peer review, editors can maintain conservatism or facilitate nuanced advancement.

This article brings to the fore some empirically grounded reasons why social-epistemic diversity can be beneficial as a countermeasure against known biases. Peer review has yet another – hitherto underappreciated – role in the larger socio-epistemic process: It can serve as a mechanism for recognizing and building expertise. Especially for young scholars, writing referee reports has the potential to develop their profile as an expert and can serve as a form of acknowledgment.

1.1 Peer review as interpersonal social epistemology in two regards

Longino's social epistemology strives for an understanding of knowledge as objective and contextual by emphasizing the relational

nature of generating scientific knowledge as part of a social collaborative and institutionalized enterprise. "… [T]he objectivity of scientific inquiry is a consequence of this inquiry's being a social, and not an individual enterprise" (Longino 1990: 67). As part of a social communal process the function of peer review is "to make sure that […] the authors have interpreted the data in a way that is free of their subjective preferences" by "bring[ing] to bear another point of view on the phenomena, whose expression might lead the original author(s) to revise the way they think about and present their observations and conclusions" and "not just to check that the data seem right and the conclusions well-reasoned" (Longino 1990: 68pp). However, she also emphasizes that peer criticism continues after publication as published work is only gradually accepted as scientific knowledge through iterative scrutiny within a cooperative dimension (e.g., through replication). We can therefore understand peer review in its two-dimensional social epistemology: first, as an interpersonal process in which the quality of a manuscript is negotiated within the triad of author, referee, and editor, whose assessments and responses collectively determine whether the manuscript is accepted by the targeted journal; and second, as a process within the scientific community aimed at reinforcing disciplinary trust and communicating that trust both within the discipline and to the general public (Jukola 2017: 125).

One of the known drawbacks of peer review is that "its confidentiality and privacy make it the vehicle for the entrenchment of established views" and Longino calls for acknowledging "effective criticism that advances understanding should be as valuable as original research that opens up new domains for understanding" while "pedestrian, routine criticism should be valued comparably to pedestrian and routine 'original research'" (Longino 1990: 76pp). Additionally, according to Biagioli (2002), by the 19th century, peer review had transitioned from being a disciplinary technique to becoming a discipline that constituted academic fields. "[P]eer review […] has been elevated into a 'principle' – a unifying principle for a remarkably fragmented field" (ibid.: 34).

He compares peer review with the market, a space shaped by "intelligent judgement, but also by low expertise, fraud, haste, and excessive conservatism" (ibid.: 35). At the same time, peer review aims to achieve the self-governance of science through rational judgments and decisions.

Jukola's (2017: 125pp) inquiry into biases in peer review also distinguishes between the intersubjective level and the institutional level of the peer review process when evaluating this scientific practice. She agrees with Lee et al. (2013: 5) that although the epistemic criteria applied by referees when evaluating pre-publication manuscripts may be labeled as "soundness, originality, scope, and clarity" the connotations of these criteria in the context of refereeing depend at least partially on disciplinary factors (Jukola 2017: 128).

1.2 Known biases and positionality of peer review – referees' role

Apart from the frequently mentioned task in the available literature that peer review should identify manuscripts deserving of publication, a second repeatedly noted task is to enhance the overall quality of research through the critical scrutiny applied in the peer review process. Reducing individual level biases, meaning biases in peer review arising from referees' or editors' "undue preferences", refers to and requires "institution-level counteractions" (Jukola 2017: 126). Here, Jukola focuses on community level mechanisms that effect the outcomes of journal peer review (ibid.: 127). She argues that the institutional level of peer review can "correct or amplify individual level biases" (ibid.: 128). The term *positionality* is used in a broad sense, as introduced above. It indicates that scholars are aware of, can reflect on, and abstract from their standpoint, situated at a particular institution, or from their personal experience when reviewing the work of others.

Regarding peer disagreement of referee reports, Jukola (2017) and Lee et al. (2013) contrast with Heesen and Bright (2021), who argue that this is a sign of dysfunctionality in peer review. Heesen and

Bright, after a comprehensive evaluation of pre-publication peer review, argue convincingly for several improvements (e.g., the use of preprint platforms and open peer review OPR), but ultimately propose to abolish pre-publication peer review altogether. Nevertheless, they acknowledge that pre-publication peer review counteracts the 'Mathew effect'. In particular, they acknowledge that the current peer review system allows lesser-known scholars to enter academic discourse once their work has been approved through a double-blind review process. However, without pre-publication peer review, one would have to rely on a network of readers active on open science platforms to build a reputation as established in fields such as physics and psychology. On the other hand, the selection of referees involves a lot of editorial decisions, because a lack of epistemic diversity within these editorial processes can affect the selection of a small set of referees which "can lead to knowledge homogenization, a perpetuation of the 'Matthew effect' in scholarly research" as also Tennant and Ross-Hellauer (2020: 3) point out.

How to distinguish between warranted and unwarranted peer review, as part of the editorial tasks, where subjective bias or personal interest of individual peer reviewers can be mitigated by editorial decisions is described in-depth and great detail by Hirschauer (2010). Disagreement can thus be due to bias but also as be a part of the critical scrutiny that makes science trustworthy as a social enterprise. While Jukola (2017) argues that the critical scrutiny inherent in the back-and-forth review process poses a potential danger to rejecting manuscripts that require revision, which can disadvantage early career scholars, Timmermans and Tavory (2022a) and Timmermans and Tavory (2022b) argue that especially the revising of manuscripts brings out the hidden quality of many manuscripts.

1.3 Conservatism and advancing a field

Another bias that could affect the plurality of epistemic stances and hinder the advancement of a discipline or field arises when

peer review is conservative against unorthodox work or biased against novel approaches. If only a few experts referee for a specific field, thereby controlling the dissemination of accepted submissions, there is a risk that the resulting work becomes more uniform than it would if the reviewers were more epistemically diverse. Content-based and confirmation biases might be increased in this way, as referees might favor manuscripts aligning with their own theoretical perspectives (see Jukola 2017; Heesen and Bright 2021).

In the social sciences, however, Timmermans and Tavory (2022a) point out that reviewers can shape the publication process actively. Reviewers may initiate a new round of data analysis by focusing their comments on three areas: the fit between observations and theoretical claims, the plausibility of the theoretical framing or explanation compared to other possible explanations, and the issue or relevance or the contribution to scholarships.

So, results are mixed regarding the topic of bias and innovation versus conservatism. The question remains of how to advance a field while addressing all these biases.

There is still a call for more empirical research on the publication process (e.g., Heesen and Bright 2021; Tennant and Ross-Hellauer 2020), but beyond that, these studies provide a roadmap for future investigations into the dynamics and processes of peer review. Tennant and Ross-Hellauer (2020) are concerned with questions that are difficult to address with quantitative methods alone. Qualitative approaches are also of relevance here, but in the end both methods will have to work together to address the functionality and quality of peer review to answer how peer review impacts scientific discourse. To address social and epistemic impacts of peer review we need to better understand the influence of peer review on innovation versus conservatism and how epistemic diversity affects the peer review process.

These questions address the tension between facilitating scientific progress and ensuring quality control in this kind of process. They also underscore the continuing relevance of prepublication peer review. In light of high editorial rejection rates

and the anticipated increase in unfiltered manuscript submissions – particularly with the rise of AI-generated content – studying epistemic criteria as a human collaborative practice is crucial during this transitional stage.

2. Methods

This qualitative study relies on empirical data from interviews with three focus groups (FG) and from three interviews with individual scholars. The participants were philosophers of science with different levels of referee and/or editing experience. The study analyzes descriptions of participants' experiences as referees for journals or in their role as editors. It assembles similarities and differences in the application of epistemic criteria to manuscripts and offers first- and third-person perspectives of involved scholars in peer review.

Research Design: This sub-study, which focuses on assessing epistemic criteria for peer review and the role of self-correction in the peer review process, is part of the larger JUKNOW project[3], studying collaborative scientific reasoning in the field of Judgment Decision Making (JDM). The sub-study with philosophers was initiated after analyzing the initial scientists' data, which showed that the comparison within the datasets (between social psychology, cognitive psychology, and experimental behavioral economics) revealed more similarities than differences. Consequently, the motivation was to recruit participants from a different discipline using different methods to enable more fruitful comparisons between participants from science and the philosophy of science, or between scholars who primarily use experimentation and those who do not. All interviews were conducted by the author. For this paper, only the philosophers' data is taken into account.

Participants for the sub-study and sampling strategy: The participants come from the philosophy of science; with some also referring to experiences in History and Philosophy of Science (HPS) or analytic philosophy more generally. None of the participants use qualitative

methods or engage in experimental philosophy. Convenience and purposeful sampling resulted in three differently composed focus groups and three individual interviews (see Table 9.1). Participants were invited to reflect on their active role as referees. Moreover, they shared their own experiences with receiving referee reports. One of the participants was part of two different focus groups, first more in their role as reviewer, second in their role as editor. Participants were either postdocs, untenured, or tenured faculty.[4] All participants had experience in publishing and had either refereed and/or edited for philosophy of science journals, books, or conferences. The following presentation of the analysis distinguishes between three levels of referee experience, based on participants' self-reports: *less experienced referees,* who describe themselves as beginners or report having reviewed fewer than ten times; *very experienced referees*, who report many years of refereeing for journals and also serving as editors, e.g., up to three manuscript a week (P3), or regularly about ten manuscripts a year for many years (P10); and *experienced referees*, who fall between these two groups.

Method of analysis: The first question was: "When you take the active role as a reviewer, what do you look at in epistemic terms to assess the quality of an article?" Interviews were transcribed by using noScribe[5], manually checked, and anonymized. Recurring themes were identified by creating mind maps and writing

Table 9.1: Participants.

		Absolut	**in %**
Gender	Male	7	63,6
	Female	4	36,4
Permanent	Yes	5	45,5
	No	6	54,5
PhD/Europe	Yes	4	36,4
	No	7	63,6
Career	Postdoc	4	36,4
	Prof	7	63,6
Sum		**11**	**100,0**

memos using narrative analysis. Then the interviews were coded with 15 codes[6] and analyzed according to experience levels and code co-occurrence.

3. Findings and integration of insights from literature

The study[7] presented here offers targeted insights based on qualitative data, derived from reflections on oral and informal professional discourse in the peer review process, as well as descriptions of epistemic criteria applied by active reviewers and perspectives from authors receiving reviews, in the philosophy of science.

Less experienced referees expressed bewilderment that there is neither formal training nor much discussion on quality assessment through peer review during academic education. P1: "[…] the thing is, nobody teaches us how to review. It's just not part of our education, which is hysterical [incredible], since it is kind of the quality measure, right?"

One general agreement was that an aim of reviewing (either by editors or reviewers) is to be selective, as peer-reviewed papers are implicitly recommended to readers. Peer reviewing, as well as the practice of doing philosophy itself (P2), is described as a form of self-correction within philosophy, where peers challenge each other's perspectives and where rejecting substandard papers helps maintain quality standards. From the reviewers' perspective, there is also agreement that "revise and resubmit" is the most common verdict. Even when clear criteria for rejecting a manuscript exist, reviewers sometimes choose to reject a manuscript themselves. However, it is not uncommon for them to leave this decision to the editors. Overall, the importance of providing constructive feedback rather than simply rejecting papers is repeatedly mentioned.

3.1 Epistemic criteria and quality control

In a nutshell, referees in the philosophy of science look for: a "novel, interesting thesis, clearly stated and supported by

a cogent argument" (P3). In addition, clarity in writing was repeatedly emphasized, along with the importance of ensuring that the argument is well informed by the literature and adds a new perspective to it.

3.1.1 Clarity in writing and the quality of language

Clarity as a quality criterion is discussed in all focus groups and mentioned in all individual interviews. An experienced referee states: P4: "[…] I would say that the language has to be clear, right? It doesn't have to be pretty or, I don't know, British or American idiomatic English. I actually feel non-native speakers —, there's often a way in which […] they write very clear because they're sort of very correct." Another experienced referee identifies the criteria they look for when reviewing a manuscript as including the statement of a novel or "bold" claim and clarity of writing.

> P5: But generally, when I review, I see these aspects: First of all, the paper should (.) have a claim. […] you see some papers which are just the review of the literature saying that this says that, that says that. [@(.)@ all chuckling quietly] There is no claim. […] And another thing, would be (.) the writing of the paper, […]. For me, usually clarity is the most important factor. It's important to be concise, but clarity is more important. […] If any criteria somehow reduce the clarity level of the paper, that's not good. […]

While most participants agree that anything not promoting the argument itself, including language, counts against clarity, others indicate that the ability of an author to express themselves is part of demonstrating their expertise.

> P6: I do find whatever intuitive strategy I use to evaluate a paper is disrupted by non-native English or just (.) English with problems. @(.)@ And so I find that I have to do a lot of explicit compensatory reasoning because I think a lot of it is just really turns of phrase that can indicate to me that intuitively someone really

> knows what's going on here. And those tend to be missing, and so you have to bring in a lot of reflection, and it becomes a much more difficult job. They can come across as an insider if they use a phrase a certain way, and that gives a lot of information that like, 'Oh, this person probably knows about this'. I'm just disposed to be more charitable in my interpretation as compared to someone who didn't use [a phrase […] that gives away whether you're an insider or an outsider].

While some appreciate linguistic competence, others criticize the use of overly complex language.

> P5: Usually, we have two kinds of authors, native English speakers [or] writers, […] they will write in an elegant way that it would become complex and that's not good. […] So, there are different kinds of problems. They write in a complex way, native speakers, and it's not good for philosophy because philosophy is complex in itself, so we should write in the most straightforward, clear way. (.) And for non-native speakers usually the problem is with this artifact of translation, grammar, choice of words and so on.

All of these aspects, technicality of language, use of terms or concepts that mark an author as an insider or outsider, in addition to personal style and flow of writing style, can be subsumed under what promotes or impedes the clarity of a manuscript. It also illustrates why both understanding and writing philosophy papers are part of the training and often take years to master.

Aspects concerning artifacts in translation and non-native English authors are repeatedly mentioned. Additionally, a very experienced scholar reflects that a lack of clarity can result either from limited proficiency in English or from unclear thinking.

> P2: [E]specially since we are all increasingly encouraged to consider that the people may not be native speakers of English. […] Sometimes I find it a bit difficult because you don't always know exactly whether […] imprecise formulations are simply linguistic in nature, i.e., the person is simply unable to express themselves

> very well, or whether it is simply thought out in an imprecise way. And I actually think that even people who are not [native speakers] have to — (.) yes (.) that is, so to speak, the core business of philosophy, to be terminologically precise. And if it's somehow ambiguous, well, then I play the usual game we play. Then I say, does he mean this or that or that? And if that, then it doesn't convince me, and if that, then I would handle it differently.

However, clarity in writing and good English are not sufficient, if the contribution is missing, e.g., P1: "I've read published papers by native speakers of English, which are grammatically and spellwise impeccable and have no cognitive content. It's impenetrable to what they are saying is utterly vacuous." [others agreeing] For FG3 using a particular type of reasoning well e.g., Bayesian expressions or Heideggerian language when at the same time missing to present a new perspective how to think about a philosophical problem or issue differently, is a reason for rejection.

Promoting clarity: clear language is paramount, whether from non-native or native English speakers. The writing does not have to be stylistically beautiful, but it must be clear (P1, P4, P7). It's the business of philosophy to be clear and precise (P2) and to use unambiguous language. Clarity is most important, alongside having a claim (P5).

Non-native English authors face the disadvantage of "missing turns of phrase" that can indicate whether an author is an insider or not (P6). Their writing can be unclear due to translation artifacts (P5) or due to untidy grammar or spelling mistakes (P7). Conversely, non-native English, is perceived as promoting philosophical argument with clear and precise language (P4), without style elements native speakers might add. A case could even be made that, due to the dominance of English as a working language, "unique idiosyncrasies and 'quirks' from other languages" can benefit philosophical discourse, provided the special meaning of expressions, which exist only in those languages, is explained to the English-speaking audience (P1).

The different layers of clarity were debated across different contexts during the interviews, as summarized in Table 9.2.

Table 9.2: Descriptions of clarity in philosophical writing.

General		**Non-native English Speakers**	
Promoting clarity	*Ways of being unclear*		*Potential advantages*
Clarity is more important than elegance	Complex language that adds no value	Translation artifacts, awkward grammar, spelling mistakes	Idiosyncratic or cross-linguistic expressions can enrich discourse
Less stylistic clutter, direct focus on argument	Ambiguity of language	Lack of idiomatic expressions	
Language should serve the argument, not obscure it			

3.1.2 A contribution in philosophical manuscripts or philosophy with problems

The positioning of one's work within the existing literature is closely tied to how its contribution is evaluated. Novelty is described as an epistemic criterion – meaning that something has not been said before, "even if it's a subtle novelty" (P4), it presents a new perspective, prompting the reader to reconsider the subject and gain new insights (FG1, FG3, P8, P3, P2).

> P4: […] I would say yeah, a paper has to […] move the needle somehow. […] But, yeah, so that's one thing I look for. […] after [reading] a paper, I'm like, 'Yeah, I knew that'. I didn't learn anything new, I didn't update in any way, I didn't change my mind, […] I am going to reject it. So, it has to change my way of thinking about it. And I think it's fundamentally the same thing as P1 said. (.) Yeah, it should give me a new way to think about a problem. And it can be an old problem, but it has to […] make me change my view, even a little bit.

There is agreement between the philosophy of science participants (in focus groups and individual interviews) that the novelty claim entails something worth thinking about and to show with a cogent argument how to think about it differently (FG3). A manuscript should have a clear, cogently stated thesis, interesting because it adds to the literature. Some participants show subjective awareness, "if it is interesting for me, it might interest others" (P6, P3). The argument should be well embedded in the literature and appropriate for and aligned with the standards of a high-tier journal, ensuring that it is not overly simplistic. For philosophy manuscripts, it is beneficial to state the claim and outline how the argument will proceed early in the text, preferably in the introduction, but it's not a necessary condition (P3).

However, when evaluating texts in an interdisciplinary setting, these criteria that participants agreed upon for assessing philosophy do not necessarily apply.

> P9: I remember thinking when I was reviewing the conference abstracts, […] I'm not really interested in whether or not people are claiming to fill some sort of gap in the literature. And I'm not even really interested in guessing out whether or not they're making an argument and evaluating the argument because that's also something that philosophers do, and that's not the kind of thing that really generalizes to a lot of other fields.

However, aside from these obvious criteria, FG1 discusses and agrees that quality assessment is not as straight forward as one might think, especially when considering its transfer into measurement, as this postdoc reflects:

> P9: […] the idea that what you're supposed to do for peer review [..] is to assess quality just seems like utterly and fundamentally mysterious to me. […] I can understand in terms of like, do you think that the person has a competent grasp of literature, is the argument that they're making coherent, is it self-undermining, do they contradict themselves, you know, (.) basic philosophical criteria for making an argument. Or, do you think this idea is (.)

> interesting enough to be worth publishing(?), like, those all to me seem to be reasonable grounds for assessment […] because I have this experience when I read, I think I get what it looks like to encounter a journal article that you think is a really, really good as opposed to something that you think is kind of mediocre, but I don't feel, I have a sense of what that means apart from, you know, being able to decompose it into (.) whether or not you think the argument is good, or valid, or coherent […]. […] ever, since I started reading about, you know, like [specification] theoretical and empirical work on peer review I've always been mystified by this talk about quality @(.)@ and @I'm not sure@ that I ever got any clarity about that, but it became a big point of discussion for these people who are advocating for automating peer review assessments (.).

Editors and reviewers report requesting authors to remove sections that do not significantly contribute to the overall argument of the paper, P10: "Sometimes, (.) one kind of change that I don't hesitate to make is, 'I think you could drop this section, you don't need this to establish your argument', because many times you see submissions where the author has felt some need to discuss almost everything connected to their topic, but really, it's not necessary".

However, participants report different practices on how to handle a manuscript that argues for something believed to be mistaken. P10 in the role of an experienced editor states: "It happens often that I see an interesting argument, which if, you know, if taken as correct would be a contribution to the literature. I think, I know how to respond to it, but I don't think that that necessarily has to mean it's not publishable." Others in their role as referees (P11) engage with what they think is mistaken and engage with the author in a critical dialogue, P11: "[…] I say, I don't know, maybe he could answer this objection. I don't see how […], I usually ask when the paper is well written, interesting topic [and] I see an objection and so I put it forward, but I say, 'Maybe you can answer to this objection', and ask for a revise and resubmit."

Jukola (2017) has pointed out that this dialogue might be short-lived, if editors subsequently decide to reject the paper based on

the referee report. However, both the scientific community and editorial boards should engage in broader discussions about the extent to which critical scrutiny benefits the discipline.

Apart from looking for clearly articulated and novel perspectives, a more experienced referee lists a couple of criteria for a manuscript to pass and gives reasons to reject a manuscript:

> P4: I definitely (.) get (.) less friendly towards a paper when I feel the author is trying to (.) sell an idea as new that's not new, that clearly exists in the literature, […] and they're not citing it, or sort of dismissively citing it. […] So, and […] I often get papers that are kind of mathematical, and I definitely am a bit nitpicky about getting the math right. I really am getting very frustrated if people throw around formulas, but they don't explain them, or they're not well-defined, or they clearly […] don't know what they're talking about, […] I always check the proofs, but also just getting the formulas right. I feel if people make just math mistakes, that's definitely a bit of a red flag for me, especially if it's important to an argument.

When scientists venture into philosophy journals, philosophical aspects might not always be well-executed and the referee has to decide how to integrate different standards from the involved disciplines.

Otherwise, articles with a strong grasp of literature or the use of mathematical formulas may seem impressive for a philosophy journal, but would not be as impressive in the author's home discipline journal (P10, P6).

> P6: when you're reviewing work that is somewhere in between a science and philosophy. Some of the papers that I've reviewed have been in psychology journals and they could have been in the philosophy journal and vice versa and (.) then you're dealing with [at least] two separate standards. You could evaluate the paper for a psychological paper, for a philosophical paper and for something else. @(.)@ And […], some of the normal considerations drop out, like impressiveness. You can have someone who's clearly a philosopher writing the paper, but has a very impressive

> grip on the empirical literature and sometimes you see reviewers being swayed by that and accepting papers just on the basis of how empirically impressive they are. […] I think that can be a mistake. (.) And if we were evaluating it by empirical standards, it just wouldn't be a contribution to the scientific literature, it would be a nice digestion of it by a philosopher. For evaluating it on its philosophical merits, it (.) might be a bit too thin. This is a common issue, I think, in reviewing for philosophy of psychology, […] venues, you have a problem between the cracks. But that's where things get even more complicated because you're dealing with different sets of standards.

For a manuscript to be accepted in a philosophy of science journal, it has to be about science in one way or another. Not "someone who has a hobby and wants to accommodate his theory of the world here. Or, that's one case and the other case is it's clearly not the philosophy of science. […] There is a relatively high percentage of manuscripts that come up regularly where it is obvious that this is not someone who comes from academic philosophy." (P2)

Less experienced scholars tend to be more willing to put more effort in their referee reports to help the paper progress than senior scholars. One reason for this could be that those experienced scholars strictly differentiate between their educational role of helping their graduate students improve their papers and their approach to other manuscripts.

On the other hand, some highly experienced scholars in philosophy (P8, P2, P10) also mention that they apply the same criteria to anonymous peer review as they do when evaluating their graduate students' papers. In general, the effort put into a referee report depends subjectively on the available time, professional interest in the manuscript, and the readability or enjoyability of the text.

3.2 Reviewers and editors reflecting on biases and positionality

When participants discuss possible biases in their role as both referees and editors, they often reflect on their own positionality in

undertaking this academic service. Meta-reflections about peer review regularly shift between their individual experiences – as referees, editors, or authors – and their awareness of being part of an evaluative practice, while also expanding to consider the community aspect of peer review as a means of quality control for a discipline. Furthermore, it is repeatedly mentioned that the academic service of providing peer review is uncompensated.

Without prompting, participants discussed various biases and demonstrated a heightened awareness of the need to mitigate potential biases (such as confirmation bias, or publication bias). In the literature, these concerns are associated with the anonymity or double-blind nature of peer review, conditions that are unlikely in an interactive academic setting of criticism or scrutiny, such as those found at conferences or in more informal collegial debates (see Jukola 2017).

Awareness of the power dynamics between the reviewer and author leads to reflections on epistemic criteria "orthogonal to" the generally accepted quality criteria.

> P9: Yeah. I think it's come up in a couple of comments already, but to be in the position of being a reviewer and to be in the position of being reviewed is to be in a […] relationship that's characterized by certain power dynamics. […] So, another difficulty that I found in reviewing the conference proposals was I have a particular perspective about the kinds of things that I would like to see at this conference. And I went back and forth for a long time about whether or not it's appropriate to bring that to bear in the kinds of things that I recommend for acceptance or the kinds of things that I don't recommend for acceptance. And that's orthogonal to questions about quality or the validity of arguments or whether or not there's a claim or any of the other kinds of criteria that we were deconstructing a quality assessment into.

This aligns with earlier reflections (P3, P6) regarding what a referee finds interesting – either personally or due to knowledge of the arguments within a specific debate.

The topic of biases is brought up in both positive and negative ways: In a positive sense, the bias stemming from the suspicion of guessing the author of a manuscript one has agreed to review, or from a personal interest in the topic of a well-written manuscript, results in a careful effort to help improve the manuscript.

> P7: So, in the case where I was suspicious that I knew the author, it's not someone that I know really well. It's someone that I've met a few times at conferences and have a positive impression of. (2) I sometimes wonder whether that influenced the way that I conducted the reviewing process. I certainly think I was maybe more patient (.) than I might otherwise have been. So having in mind, like, I like this person, let's think about what's written a little bit more carefully and […] read it in a slightly more charitable way […], you know, take a little bit more trouble to suggest how to improve rather than being a little bit harsher as intended. But the effect at the time, it felt so slight, but it seemed maybe not worth (.) troubling too much about.

The same participant also stated to be more engaged when the paper is interesting to them and well-written.

> P7: […] how much time I spend on a review varies based on how much I @enjoy the paper@. I'm going to be totally honest. […]. I spend more time if it's interesting, if I think it's good, I think it's well-written. Yeah, and much less time if […] my first impressions are more negative. So, there's engagement with the topic, like how personally interesting I find the topic. And then also there's considerations of just at first glance, how well the paper engages with it. […] But if it's an interesting paper that is well-written, nonetheless has some room for improvement, there's like a sweet spot, which will kind of captivate me and then I'll accidentally spend a really long time on it.

When a less experienced referee asked an experienced one whether knowing the author would influence the way they review a manuscript, the latter answered:

> P4: Erm, I'd like to say no, but I'm not sure, but I'd like to say no. And also, it's not that I know the author personally. I just know who they are. […] Right, I think if it was a friend of mine, I would definitely say no. […] But in these other cases, I had strong suspicions just based on knowing people's work and not based on knowing them personally.

The ethos for being a referee is explicated by P3: "Refereeing is done anonymously and with the obligation (.) that you not draw on the material that you read. So, I *have* to forget and it's very easy to do."

From the editor's perspective, actively counteracting being biased against non-native English speakers was explicitly mentioned:

> P10: […] some papers are written by non-native English speakers, submitted in English, and as an editor you have to watch out that referees are not being dismissive just because of grammatical errors or poor word choice, which is hard to overcome. I mean, it's getting easier to overcome because of ChatGPT and Google Translate and things like that, but (.) it's important both as a referee and as an editor to try to be super fair and make sure that the refereeing is done on the basis of content, not these kind of flaws with writing that are almost inevitable when it's not your native language.

Recently, implemented editorial practice to share referee reports among the other referees is discussed in FG2 and FG3 and is commonly appreciated, e.g., P10: "it can provide of just making you see how a paper has been criticized from a completely different person's perspective can be very useful".

From the editor's perspective, choosing the right referees is crucial. Participating editors shared pragmatic thoughts about that: reviewers have to return the report within a reasonable period of time and provide helpful assessment for the editor who decides on major revision or rejection. In all individual interviews participants highlighted the problem of finding adequate referees who can return the manuscript promptly, noting that they either use

their own network of scholars or refer to assistance from other members of the editorial board to help with this task.

Editors must balance their decisions between their assessment of the quality of the review report, including evaluating possible bias or factors of positionality of the reviewers, their own assessment of the manuscript – if they choose to read it – and sometimes also the number of potentially publishable submissions a journal receives and requires. They interfere if: "you can tell that [those] referees are trying to (.) tell the author to rewrite the paper the way that they would have written it if they were the author and had the idea. And that can (.) be, I think, […] unfair to the authors and also lead to a paper that is skewed in a direction the author didn't want to go." (P10).

When asked how they would react in this case, they said:

> P10: I would tell the author, 'Please definitely respond and make changes in respect to this comment or this other comment. However, you can treat the rest of what referee number two says as optional advice. […]' And then I wouldn't resend the paper to that same referee if I'm telling an author, you know, 'Take what you want from this referee report, but don't feel obliged', then of course, I'm not going to expose them to the ire of a referee whose advice hasn't been taken. @(.)@

Although many of the desk rejections happen at the editorial level, the role of the referee is to protect the reader: "I'm protecting readers. […] If I as a referee, who has an obligation to read the paper, get so fed up that I don't want to read it anymore, who's going to read this(!?), right? I mean, I do value clarity in writing and it's got to be done, it's got to communicate well." (P3)

Another challenge posed to referees is the requirement to competently assess a manuscript to decide whether a contribution is *a benign, minor addition hardly worth publishing*, *a small but subtle contribution, a major contribution*, or *a plausibly argued fundamental novelty*. Undoubtedly, these judgments are difficult to make and require a self-reflective attitude regarding one's own

positionality in the field, self-interest, along with an extensive overview of the literature discourse. In this context, P3 points to the problem of a referee's caution, noting that it is often safer to label a contribution that is a possible 'fundamental novelty', as "a foolish paper, where it's very unlikely that everyone else has been mistaken" (P3). In short, contribution, bias, and novelty appear to be in tension within the subjective assessment of an individual referee. We will return to this point in the next section.

Ways to react to conflict of interest that were mentioned by the participants are: a) to refuse to referee a paper if it's too close to one's personal interest (P4, P3); b) to *sign the referee report* and thereby to unblind the referee (P3).

The bias towards publishing a manuscript rather than rejecting it is also discussed in relation to the pressure to publish. This bias is mentioned as a secondary theme in the assessed data but arises twice when participants reflect on the quality of referee reports and the role of referees and editors in improving manuscripts that require substantial revisions, particularly to fill a journal, especially in the context of journals in their early stages. P8: "Of course, they reject 90% of the bullshit, but in the end the 10% are published. This is exactly the same in our journals. You insist, you send it to a first journal, they say, 'Maybe not', you send it to a second one and they say 'Yes'. They [journals in the beginning phase] need the paper, in fact, […] so that's it." Or by P10: "when I was editing. I read the referee reports and (.) intended to be more favorable to the author. So, if there was a positive and a so-so report, I would give the author a chance to revise and resubmit. But that's also because it was a new journal starting up that I was editing." Compared to P10: "journals which simply have huge (.) numbers of submissions, so, the acceptance rate is exceptionally low and if you have even one referee who's like, 'I'm not sure', sometimes an editor makes the choice, […] [to] only publish ones that have two very positive verdicts."

Especially Heesen and Bright (2021) highlight the unclear benefit of pre-publication peer review, given the various biases inherent in this academic service. Although in the present interview

study participants discussed their individual biases – without naming them as such – none questioned the existence of pre-publication peer review. However, its functionality and potential improvements were debated creatively across various aspects.

If editors mediate between authors and reviewers, they aim to protect against injustice or unfair treatment. This is acknowledged from the authors' perspective, where they may seek editorial assistance when receiving reviews, which involves mitigating harm or avoiding unfair treatment of authors by reviewers.

> P9: I think there's some really predictable […] forms of responses that people give in reviewer reports. There is all this talk about the dreaded reviewer, […] (.) the basically uncharitable reading of your work such that they mistake key components of your argument or they just misstate or misunderstand key claims that you're making and I think responding to those is pretty straightforward and sometimes the kind of thing for which you need an editor to intervene but sometimes not.

How much influence should reviewers be allowed to have? Due to social acceptability bias, none of the participants identified with the "dreaded reviewer"; consequently, descriptions and reflections of negative experiences are only reported from the receiving end. P9: "Maybe this is giving reviewers too much influence over the final product of the work, when the reviewers are just a very small sub-selection of people who have thoughts about the work. So that can create certain kinds of influence that we might think on the whole, collectively, could be pretty problematic."

The point of interference was picked up by participants of all experience levels. P6 (experienced): "But I think it is important to avoid some kind of imperialism on the part of the reviewer and to clearly flag that, so that the author isn't in an awkward position if they don't want to write the paper the reviewer thinks would address the problem. And so that the editor doesn't have to step in and look for a new reviewer, which is an intensive process."

Another experienced reviewer (P5) argued that, because the positive review of a manuscript is an implicit recommendation,

it is the responsibility of reviewers to be selective. P5: "[…] now after about three or four years of reviewing, I'm more inclined to reject because my experience in these years is that the second submission would not be substantially different. It just takes the time of me and the author him- or herself. […] So now I'm more inclined […] to reject".

Interference from reviewers with the interpretation is discussed in the literature in two ways: as legitimate (see Timmermans and Tavory 2022a) and as both legitimate and illegitimate (see Hirschauer 2010). In the first instance, the paper profits significantly from the scrutiny and questions posed by the critical reader, leading to multiple revisions that may transform the manuscript into a totally different, but at the same time greatly improved and consensually accepted paper. Conversely, unfair and misaligned criticism, unhelpful or impolite personal comments, or aspects related to positionality show power dynamics among the involved actors.

The participating philosophers, on the other hand, describe self-correction of their field as an inherent aspect of their work, part of their routine practices of identifying and addressing issues in the arguments of others to enhance them. In interdisciplinary settings, however, differing epistemic cultures of criticism can have a negative effect on how a discipline is perceived, as Lamont (2009) has diagnosed in her book "How Professors Think". An unhelpfully harsh review style of philosophers when evaluating academic excellence of their peers and others may lead to the downgrading of their peers' chances for securing funding and having manuscripts accepted (see Jukola 2017).

In contrast, especially FG3 discussed the joy of this academic practice. They agreed that being asked to review for the first time is perceived as an initiation and acknowledgment of one's entrance into academia – while the lack of compensation often becomes a joking point.

> P7: I actually felt like it was quite a privilege because my Ph.D. was funded for the first three years, //P4 Yeah.// […] and when I got asked to review a paper, I felt like, 'Oh, privilege', like 'Ah, I'm

> a real @academic now@'(!). <P1: Welcome to the world of unpaid labor in academia, yes(!), what a privilege.> [all laughing] No, but I was so happy about it. I was like, 'They're trusting me(?)'. <P1: @(.)@ I know. No, they're desperate. They need people to review for them, for free.> [all laughing] I spent a long time on it and thoroughly enjoyed it. And then gradually it gets —, kind of, you get a bit more grumpy about it. […] <P1: Same for me. I was also so excited the first time to be even asked my opinion, //P4 Same, yeah.// it was so exciting.> Yeah, yeah.

On the one hand, good reviewers are scarce; on the other hand, reviewers' self-interest bias is described as unhelpful in various ways. Consequently, expanding the pool of referees to include more diversity seems a logical step.

3.2.1 Time

Heesen and Bright (2021) evaluate peer review through the lens of epistemic consequentialism and argue that the high time costs peer review imposes on individual scholars could be better allocated to other academic tasks or on conducting research. Other scholars emphasize the invaluable service of calibrating reading time for others (see Hirschauer 2010; Harnad 1998).

In this study, less experienced reviewers see pre-publication peer review not only as their service to the community but also as a learning opportunity and make an effort to help to improve the manuscript. They invest significantly more time, especially when they wish to establish themselves as emerging experts. Initially, they feel honored to be entrusted with an important task for the community and are often grateful for having previously benefited from the reviews of others. Over time refereeing becomes a resource problem not only but especially for those without permanent positions.

Experienced and very experienced referees report that they rarely spend more than 24 to 48 hours for assessing a manuscript and writing their report. They justify this limited time by arguing that quality assessment is straightforward. At the same time,

this group also includes individuals who state that they sometimes need only one to three hours to assess the quality of a manuscript. However, despite the speed of their reviewing, the overall process might still extend up to two months.

One of the very experienced referees described the process of a fast turnaround. P3: "And I'll probably start reading it on the bus home, because I want to get a very fast turnaround. So, if it's the sort of paper that I can deal with quickly, I'll have the report written after dinner and sent off. If it's a paper that requires a lot more thought and a lot more work, it'll maybe take another day or so."

One focus group (FG3) discussed the time taken for refereeing in more detail, uncovering significant differences between less experienced and more experienced referees. They reflected on how "general thinking time" is part of the methodology in philosophy, noting that what is wrong with a paper might reveal itself when one is not actively thinking about it.

> //I: Can you say what a long time is?// P7: Good question. (.) Really hard to say. Like, I guess I'd finish it within a week. I probably wouldn't get very much else done during that week. So, say like two to three working days, possibly. (.) Hard to say, yeah. I'd say like, yeah, three, between two and three working days, yeah. Where not much else happens. (.) Which is quite a long time. (.)

The same participant elaborated on how the timing changed over the course of refereeing experience. P7: "I really can't remember the first couple of times, it was maybe part way through my Ph.D., I was asked to review a couple of things. I remember it taking a very long time, but I sort of thought that was fine because everything takes a long time when you're first learning how to do it."

Moreover, referees expressed their impatience with papers that appeared unfinished: P7: "And (.) then find myself being quite impatient with papers that I'm asked to review that don't look finished. //P1 Yeah// So they're the ones that I will spend much less time on and have a kind of impatient attitude towards."

In addition to that, effort put into this task is generally reduced over time. Referring to a discussion with a very experienced referee, P1 reports: "But I [recently] spoke […] to somebody in the field […] who reviews a lot, so I asked for advice. […] And that person had a very straightforward answer. They spent exactly two hours on a review, (.) by rule. […] Not multiple days, nothing else, two hours. If the paper is a good paper, I am able to see if it's a good paper in under two hours. That's it. If it takes me more than two hours to find out whether it's good or not, then it's not a good thing, simple as that."

Another very experienced referee and editor described the process in a way that might sound more familiar.

> P2: I also have to say, if I had more time, I would enjoy it more. […] I actually enjoy reviewing. And yet (.) when I think, 'Oh, great, now I have the whole day time for this', somehow, then I read and really try to get into it and consider whether I like it and what could be improved. In fact, it often becomes a great burden when you are suddenly weeks behind and still have three referee reports —, (.) or, accordingly, the editorial tasks. […] I usually do it all at the last minute, and then it all has to happen at once. I read, and then I write the report right away. The problem is that you get a deadline and I always do it when I'm reminded, so when the deadline is over and then I do it but also relatively quickly. <I: […] is that 12 hours, or 4 hours?> No, I'll be reminded after 2 months, and then it'll be, well, (2) probably more like 24 hours.

Another explanation for the time allocated, besides experience, is that faculty – whether tenured or not – cannot afford to spend more time on this sort of academic service alongside their full-time duties, which include teaching, editing, research, and supervising students.

Alongside the issue of time, it is noted that peer review is a necessary academic service that lacks extra pay or compensation and offers only indirect recognition (P8 and P2). In this context, and from an editorial perspective, P8 and P10 mentioned difficulties with insufficient reports, e.g., referees who pretend to read but

do not and subsequently write inadequate reports. FG3 discussed benefits and pitfalls of paying graduate students to conduct peer review and pondered various ideas on how to compensate for this academic service.

Both Jukola (2017) and Hirschauer (2010) allude to the varying experience levels among referees. Hirschauer writes "[…] referees only react to occasional demands for reviews, and normally have less experience as referees than as authors" (2010: 87pp). By diversifying referees in multiple directions, including levels of expertise and seniority, two objectives can be achieved: (a) the problem of reviews that are not informative enough (due to time issue) can be resolved. (b) By combining more knowledgeable referees with those who are less biased or invested in a specific direction, arguments can be evaluated for their intellectual merit and potential for future development, rather than against the backdrop of intellectual trench warfare.

So far, the existing literature on peer review has underappreciated the possibility to actively selecting referees not only based on competence but also from different experience levels. The social-epistemic aspects of this phenomenon should not be underestimated. Although editors rely on reviewers to assess papers competently and constructively and expect them to do so within a reasonable time frame, the data from this study indicates that editors do call upon early career scholars. Whether this is the case because these young scholars put in considerable effort, or whether it is just a fluke in this participant group, cannot be determined here.

3.3 Contribution – innovation – conservatism – self-correction

Peer review has shown to serve many functions, including gatekeeping and calibrating the reading time for others. It assesses the contribution of a submitted manuscript in terms of its fit with the journal and its expected interest to the targeted readership.

The peer review process also negotiates whether a contribution is considered minimal or too insubstantial to warrant publishing, a subtle but interesting or a major contribution, or a potentially fundamental novelty with the power to change the direction of a field. FG2 reflects how challenging this assessment can be.

> P11: Maybe that's also why it's difficult to decompose a paper in different (.) interest[…] and things like that because often the way you judge those dimensions are not independent. And sometimes […] one dimension is a proxy for the other one, and so sometimes […] it's well written, you end up with a huge score. Whereas if you step back, you say, 'Well, it's not so great. I mean, it's just an epicycle in the debate.' [both P6, P10 agreeing]

Thankfully, peer review is only one way to assess the novelty of an idea. Such discussions also take place at conferences, during colloquium talks, and in other formal and informal academic settings – providing opportunities to test new ideas, subject them to scrutiny, and build networks where they can develop.

Alongside the above-mentioned basic philosophical criteria for making a valid argument, participants also discussed what the quality assessment is based on, beyond the experience required to distinguish "really, really good as opposed to something that […] is […] mediocre" (P9). P3 states that "[…] all of those things are difficult to give you a simple recipe, an algorithm for. If there were (.) a recipe for that, you wouldn't need to ask referees." Referees are needed "because there're hard judgments to make. […] If people wrote their arguments in first order predicate logic, then we could have an algorithmic argument checker. But they don't. And […], for good reason, because first order predicate logic is not strong enough to do the sorts of arguments that are needed. So, it's quite difficult. And especially inductive arguments are very troublesome […]"

P8 and P1 described frustrations with referee reports that either resist innovation or state the lack of novelty of a manuscript in a single sentence without giving reasons. Both types of verdicts can

lead to the rejection of a manuscript. FG3 agreed that if a manuscript is rejected, the author should receive reasons and evidence in support of the decision. If a claim is deemed not new, reviews – understood as an inherently dialectical task – should at least provide citations to demonstrate that the argument has already been made. P8, who has a collegial relationship with the editor in charge, described his experience from the author's perspective:

> P8: The editor-in-chief […] sent me an email, 'It's very interesting, I sent it to a referee.' […] The referee […] says, 'Oh, it's a beautiful paper. We have to publish it.' So, the editor-in-chief is very excited and [...] said, 'Oh, wow, it's good.' […] And the other referee says, @'Oh, no, I don't understand a word. This is all bullshit.'@ Very negative things. So, the editor-in-chief says, 'Okay, I have to follow my referees, so I won't publish it in this state.[8]

On the other hand, P8 in their role as an editor expressed the desire to make room for new contributions, without much success. P8: "We were interested in bringing together those two approaches to make something new. And what we have in fact, we have old very well-known philosophers who say exactly the same thing as ten years ago. So, there is no progress at all. And we have young newcomers, who do something very descriptive, […] conceptual history, which is not history in itself because it's conceptual." And after experiencing a very high rejection rate, resulting in a small number of suitable submissions, P8 confessed to accepting everything.

Innovation takes time – so does self-correction:

> P3: Well, when I'm interested in philosophy of science, I'm interested in the foundational questions as to what constitutes science. We have a standard repertoire of stuff: realism, instrumentalism, @explanation, and so on@, all the way through, you know the list. So, by default, if it's in that area, I wouldn't limit it to that area. In my time, various areas have entered and become new. I think data has become a topic that never was one, but now it's clearly part of the canon, and it belongs because it's being analyzed in terms of fundamental ideas in the science. And so, yeah, again,

> this is why we don't have a secretary checking a box, it @requires@ some insight and expertise in philosophy of science to do this.

Participants described situations in which they, as referees (P11) or editors (P10), would react when they considered the content of a manuscript to be mistaken or disagreed with its argument. One approach mentioned was to request that the author address the objections raised when asking for a revision and resubmission, thereby entering into an internal dialogue. Another approach would be to allow the argument into the discussion despite believing it to be flawed (see chapter "Epistemic criteria and quality control").

Personal interest was initially mentioned in relation to biases; here, it is discussed with a focus on self-correction and innovation within the field. One of the very experienced referees and editors described the problem of the cautious reviewer which is linked to the issue of how innovative perspectives can enter a field dominated by conservatism and caution.

> P3: […] How well does [the refereeing system] actually work? It's terrible. (.) You have to understand who the referees are. When you go into a field, you discover that there's an area that's been unexplored, that people have neglected stuff, or that worse, it's been done badly, and it needs to be cleaned up […], and so you write a paper saying, 'What a shame this has been neglected', […]. Now, who are your referees going to be? You think they're going to be wise Olympian gods on high mountains, who have no interest in the field, and whose only goal is to advance the profession. No (!), your referees are exactly the people who made all the mistakes that you're criticizing. And you rely on them being able to rise above their personal interests and be able to judge you. It doesn't happen. (.) Or, it doesn't happen nearly well enough. So, you're going to get incoherent, confusing referee reports.

Especially this participant, over the course of his many decades of being a scholar in philosophy of science, goes into detail about the challenges of advancing or self-correcting the discipline. P3: "I can tell you there are a series of traditions where I can see them

living on in spite of absolute refutations." The examples come from Bayesianism: "I'm not objecting to […] probabilities in the right circumstance, but this default idea that you automatically apply probabilities to everything, it's just wrong." Nevertheless, P3 assess that if corrections happen, then slowly and more through bifurcation than through direct correction. P2 also makes the point that philosophy, as a discipline, is oriented towards analyzing and identifying weak arguments and – in this sense – self-correction within the field works well. The broader question would be whether entire directions of research might ultimately prove to be fundamentally mistaken.

It seems that peer review serves as a form of self-correction for faulty manuscripts – but not in a strict sense. Editors may decide to encourage discourse by relying on the broader community of peers to engage with and address debatable arguments after publication. This highlights an interesting point: peer review is only one factor in the broader process of self-correction and scrutiny within a field. As the previous quotes show, it is therefore not so surprising that some arguments remain in the literature despite having been refuted.

Longino (2002) already emphasized the need for a critical discursive academic culture and warns that suppressing the development of this culture can be epistemically detrimental.

4. Conclusion

This article has examined peer review as both an interpersonal process and a social-epistemic institutionalized process, grounded in a qualitative study with philosophers of science. At the interpersonal level, referees evaluate arguments, rigor, and clarity, while their judgments and reflections are influenced by considerations of epistemic criteria, biases, and positionality. At the institutional level, peer review is part of a broader process of organized criticism, where editorial decisions can influence the balance between conservatism and advancement (see Hirschauer 2010).

The findings demonstrate that peer review is best understood as a complex, dialogical practice that involves communal effort and a wide range of epistemic, pragmatic, and social considerations. While some participants criticize the peer review process in terms of bias or conservatism, none of the participants argued for abolishing it altogether. On the contrary, the complexity described throughout the findings highlights its indispensable role in self-correction, while acknowledging its conservative tension with advancing a discipline. This dynamic – between quality control and openness to novelty – remains one of the crucial challenges for the future of academic publishing. As Tennant and Ross-Hellauer (2020: 7pp) state, epistemic diversity during peer review is likely to be low in homogeneous research communities, which contributes to peer review bias and the dominance of more traditional views, with repercussions for the entire research community. There is a need for assessing and contextualizing bias in human-driven peer review to mitigate potentially harmful effects.

This article has presented some evidence that peer review functions as a form of recognition, particularly for early-career scholars, and serves as a venue for building expertise and socializing into disciplinary practices. Diversifying referees in terms of epistemic background and positionality can help mitigate bias and has the potential to counteract conservatism where innovation would be preferable for epistemic reasons. Finally, this discussion points towards the need for greater systematic reflections on the peer review process itself. Despite its central importance to science and the humanities, writing referee reports is rarely taught, openly discussed beyond anecdotal evidence, or systematically studied as a practice. To more fully understand publishing and its role in self-correction, an interdisciplinary effort is needed, drawing from the sociology of knowledge, publication process research, history of knowledge, philosophy of science, organizational research, and related fields (see Hirschauer 2010; Tennant and Ross-Hellauer 2020).

Overall, peer review remains one – though not the only – practice that ensures the academic quality of published manuscript, contributes to self-correction, and – if functioning well – fosters

innovation in academic disciplines. Its value lies precisely in its situated, diverse, and dialogical structure. Recognizing and strengthening this role is essential for peer review to continue reflecting the richness of human expert judgments and contributing meaningfully to the advancement of knowledge.

4.1 Reflexivity

As an author within the field under investigation, I maintained a reflexive awareness throughout the study. As a researcher in the philosophy of science in practice and as a referee with moderate reviewing experience, I acknowledge potential overlap with the epistemic criteria discussed. Before writing this text, I presented preliminary findings to an audience of sociologists and philosophers and incorporated their feedback. Notably, during the interviews with philosophers, participants challenged some of my preconceptions – developed during interviews with scientists and subsequent data analysis – particularly regarding self-correction and novelty. This demonstrates how empirical methods can productively interrogate the investigator's assumptions.

4.2 Limitations

Qualitative studies are not generalizable in a statistically significant way. As stated in the methods section, the findings aim to deepen the understanding of pre-publication peer review rather than provide quantitative representativeness (see Hangel and ChoGlueck 2023). Concerns about social acceptability bias are mitigated by the fact that participants willingly discussed both positive and negative aspects of their peer review experience. Despite the diversity of views and the relatively small participant group, several themes showed strong saturation across experience levels and interview types. Due to space limitations, distinctions between HPS and philosophy of science criteria, as well as topics such as blinding and de-anonymization, will be addressed in future work.

Acknowledgements

Scientists' analysis was discussed by the reading group at the center for philosophy of science and at CMU, Pittsburgh. The philosophers' analysis was presented at the LCSS Colloquium, Leibniz University Hannover. The study was funded by DFG # 500479165. The manuscript profited from the patience, improvements, and comments from the editors of this book.

Notes

[1] In philosophy, it is still rare for a publication to have more than one author. Mostly, pre-publication journal peer review is blinded although open reports, signed reviews, and other transparency practices are being discussed.

[2] Conceptual analysis of analytical philosophy versus a hermeneutic focus on interpretation of texts or meanings can serve as a methodical example here. For different theoretical epistemic styles think of coherentist versus fallibilist framing of an argument.

[3] The sub-study was conducted with the same research standards concerning study information and informed consent, see https://gepris.dfg.de/gepris/projekt/500479165.

[4] One participant, with experience in both editing and reviewing, was interviewed a few months before graduation, therefore, technically a Ph.D. student but comparable with other postdoc participants with referee experience.

[5] Wollin-Giering, S., Hoffmann, M., Höfting, J. & Ventzke, C. 2024. Automatic Transcription of English and German Qualitative Interviews. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 25. describes noScribe as best available standard for AI-assisted transcription software.

[6] The codes are: Bias_anonymity, error, HPS vs Philsci, epistemic criteria, normativity, power inc. social relations, positionality, review_active role (for all three experience levels), review_receiving, role_editor_passive, role_editor_active, self-correct vs innovate, time and/or pressure. Five of these codes were the same as those used for scientists e.g., "epistemic criteria", "time", "review receiving". Four of the existing codes were adapted to facilitate the analysis,

e.g., "reviewer_active role" was split into three to depict the different experience levels while "role_editor" was split into a passive and an active perspective. Three new codes were added: HPS, self-correction vs. innovation, and bias_anonymity, although the latter two also appear as memos in the scientist's dataset. The qualitative data software used is called *Atlas.ti*.

[7] The study presented in this article profits in many ways from another qualitative study on editorial judgments for the *Zeitschrift für Soziologie* by Hirschauer (2010). According to Hirschauer, the process of journal peer review "consists of partly written, partly oral, partly formal and partly informal professional discourse that occurs beneath the threshold of publication" (ibid.: 72). Hirschauer provides a comprehensive ethnographic account of the peer review process based on ten years of participant observation and editorial records, as he was a part of the editorial board. He reconstructs editorial decision-making as a broad, ongoing negotiation between editors and referees. Hirschauer identifies three reasons why peer review has a poor reputation: first, the procedures could be better organized (e.g., safeguards of pluralism); second, self-serving bias – when the "social dimension" of peer review is only acknowledged if the verdict is unfavorable to the author; and third, "inappropriate scientistic expectations" (ibid.: 72). Studies that try to assess quality using "validity criteria intrinsic to quantitative methodology" often find that "scientists in peer review do not work 'scientifically', without giving up the assumption that the assessment of scientific communication is a scientific measurement problem" (ibid.: 73).

[8] The paper in question was turned into two chapters of a new book.

Bibliography

Biagioli, M. (2002) "From Book Censorship to Academic Peer Review". *Emergences: Journal for the Study of Media & Composite Cultures,* 12: 11–45. https://doi.org/10.1080/1045722022000003435.

Hangel, N. & Choglueck, C. (2023) "On the pursuitworthiness of qualitative methods in empirical philosophy of science". *Studies in History and Philosophy of Science,* 98**:** 29–39. https://doi.org/10.1016/j.shpsa.2022.12.009.

Harnad, S. (1998) "Learned Inquiry and the Net: the Role of Peer Review, Peer Commentary and Copyright". *Learned Publishing,* 4: 283–292. https://doi.org/10.1087/09531519850146229.

Heesen, R. & Bright, L. K. (2021) "Is Peer Review a Good Idea?" *The British Journal for the Philosophy of Science,* 72: 635–663. https://doi.org/10.1093/bjps/axz029.

Hirschauer, S. (2010) "Editorial Judgments: A Praxeology of 'Voting' in Peer Review". *Social Studies of Science,* 40: 71–103. https://doi.org/10.1177/0306312709335405.

Jukola, S. (2017) "A Social Epistemological Inquiry into Biases in Journal Peer Review". *Perspectives on Science,* 25, 124–148. https://doi.org/10.1162/POSC_a_00237.

Lamont, M. (2009) *How Professors Think. Inside the Curious World of Academic Judgment*. Harvard, Harvard University Press.

Lee, C. J., Sugimoto, C.R., Zhang, G. & Cronin, B. (2013) "Bias in peer review". *Journal of the American Society for Information Science and Technology,* 64: 2–17. https://doi.org/10.1002/asi.22784.

Longino, H. E. (1990) *Science as Social Knowledge Values and Objectivity in Scientific Inquiry*. Princeton University Press.

Longino, H. E. (2002) *The Fate of Knowledge*. Princeton University Press.

Tennant, J. P. & Ross-Hellauer, T. (2020) "The limitations to our understanding of peer review". *Research Integrity and Peer Review,* 5. https://doi.org/10.1186/s41073-020-00092-1.

Timmermanns, S. & Tavory, I. (2022a) "Thank You, Reviewer 2: Revising as an Underappreciated Process of Data Analysis". *Sociologica,* 16: 47–58. https://doi.org/10.6092/issn.1971-8853/14665.

Timmermans, S. & Tavory, I. (2022b). *Data analysis in qualitative research Theorizing with abductive analysis*. Chicago, Chicago: University of Chicago Press.

Wollin-Giering, S., Hoffmann, M., Höfting, J. & Ventzke, C. (2024) "Automatic Transcription of English and German Qualitative Interviews". *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research,* 25.

TEIL III

Auswirkungen auf die philosophische Praxis und Lösungsvorschläge

TEIL III

Auswirkungen auf die philosophische Praxis und Lösungsvorschläge

Nicola Mößner und Klaus Erlach

Aus den vorangegangenen Untersuchungen in diesem Band ist ersichtlich geworden, dass die zunehmend digitalisierten Arbeitspraxen wissenschaftlicher Disziplinen eine Reihe von Herausforderungen und zum Teil auch schwerwiegende Probleme mit sich bringen. Zu letzteren zählt insbesondere das Wissenschaftlertracking, wie es von Bernhard Mittermaier beschrieben worden ist.

Deutlich geworden sein sollte ebenfalls, dass die aufgezeigten negativen Entwicklungen im Wissenschaftsbereich aufs Engste miteinander verflochten sind. Fehleinschätzungen wie eine vermeintlich rein quantitativ erfass- und vergleichbare Expertise im Kontext wissenschaftlicher Leistungsevaluation (vgl. z.B. Retzlaff 2022; Mößner 2024) sowie Fehlentwicklungen wie das

Zitationsvorschlag für dieses Kapitel:
Mößner, N. und Erlach, K. 2026. Teil III: Auswirkungen auf die philosophische Praxis und Lösungsvorschläge. In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum*. S. 287–294. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.l. License: CC BY 4.0

Datentracking durch Data-Analytics-Unternehmen (vgl. z.B. DFG-Ausschuss für wissenschaftliche Bibliotheken und Informationssysteme 2021; Altschaffel u.a. 2024; Siems 2022) erscheinen als Konsequenzen dieser Verquickungen.

Als besonders besorgniserregend zeichnet sich als weitere Folge der dargestellten Entwicklungen ein drohender Autonomie-Verlust der wissenschaftlichen Gemeinschaft ab (vgl. Holzer 2022). Die Wahl von Fragestellungen für wissenschaftliche Untersuchungen, ja sogar die Themensetzung für ganze Forschungsfelder stehen hier auf dem Prüfstand. Angebote der Data-Analytics-Unternehmen, Aussagen zu Forschungstrends und (vermeintlichen) Expert:innen durch einfaches Abfragen bereitgestellter Datenbanken verfügbar zu machen, stellt dabei nur einen der Aspekte dar, durch welchen die Autonomie des Wissenschaftssektors unterwandert wird (vgl. z.B. Herb 2018).

Nicht weniger problematisch ist im Arbeitsalltag der Wissenschaftler:innen die Schaffung von Abhängigkeiten von Technologien, welche mittlerweile den gesamten Forschungszyklus durchziehen (vgl. Mößner 2025). Man verdeutliche sich das am Beispiel kommerzieller Betriebssysteme und daran geknüpfter Textverarbeitungsprogramme. War es bisher möglich, diese käuflich als Eigentum zu erwerben und auf dem eigenen Rechner zu installieren, ist man nun von der Vergabe von Nutzungslizenzen abhängig. Software-Updates können in der Folge die Inkompatibilität von Programmen und mit ihrer Hilfe erstellten Arbeiten bedeuten. Will man seine Daten weiterhin nutzen und in der Community austauschen, scheint kaum noch ein Weg am Abschluss immer neuer Lizenzverträge vorbeizuführen.

Nicht minder schwer wiegen die Probleme, die durch die Nutzung sogenannter Cloud-Lösungen entstehen. Auch hier werden problematische Abhängigkeiten generiert. Einmal eingespeiste Daten sind oftmals nur schwer in andere Systeme zu migrieren. Darüber hinaus stellen sich drängende Frage zu Datenhoheit und zum Datenschutz. Je sensibler die Forschungsdaten, die in Cloud-Anwendungen abgelegt werden, desto virulenter werden diese Themen und desto größer der

Autonomie-Verlust nicht nur individuell Forschender, sondern auch der beteiligten Institutionen.

Schließlich steht durch das Datentracking der Informationsanalysten und ihre Möglichkeiten zur subtilen Reglementierung des Forschungsinputs der Nutzer:innen ihrer Datenbanken und anderer webbasierter Werkzeuge letztlich auch das Recht auf die Freiheit von Forschung und Lehre zur Disposition. Man denke im Lichte aktueller politischer Entwicklungen an Fragestellungen zum Klimawandel, zu Feminismus, Gender Studies und Diversity usw. Soll sichergestellt werden, dass die Forschung auch in diesen Bereich unabhängig und kritisch weiter vorangetrieben werden kann, ist eine Bindung an Systemanbieter, die nicht nur das Publikationsverhalten, sondern auch seitengenau das Leseverhalten von Wissenschaftler:innen nachverfolgen, analysieren und in Profilen für Drittverwender von Daten aggregieren, geradezu dystopisch.

In diesem dritten Teil *Auswirkungen auf die philosophische Praxis und Lösungsvorschläge* soll daher konkret danach gefragt werden, wie die aufgezeigte Dystopie langfristig vermieden werden kann. Im Fokus stehen dabei die folgenden Punkte:

a) Was sind zeitgemäße Wege der Verbreitung wissenschaftlicher Erkenntnis und welche Institutionen eignen sich in Ergänzung oder besser noch als Alternative zu den Datenanalysten für die Bereitstellung der entsprechenden digitalen Infrastruktur?
b) Wie kann die Autonomie der Wissenschaft bei der Bewertung ihrer wissenschaftlichen Leistungen sichergestellt werden?
c) Inwiefern muss sich auch die Wissenschaftskultur ändern, um den beschriebenen negativen Entwicklungen entgegenzuwirken, und welches Potential, welche Herausforderungen bringen Vorhaben im Bereich der Open Science mit sich?
d) Welche zentralen Themen sollten von der wissenschaftlichen Gemeinschaft am Beispiel der Philosophie angegangen werden?
e) Wie sähe eine von wissenschaftlichen Institutionen mitgestaltete Publikationsform beispielsweise aus?

Die Academia ist ein komplexes sozio-technisches Gefüge. Auch die akademische Philosophie stellt hier keine Ausnahme dar.

Oftmals machen wir uns im Arbeitsalltag nicht bewusst, wie viele Rädchen im sprichwörtlichen Getriebe der Wissenschaft ineinandergreifen, damit unsere Forschungspraxis von der Recherche über die Ergebnispräsentation und kritische Diskussion bis hin zur schlussendlichen Publikation funktioniert. Vor allem jene Einrichtungen und in ihr arbeitende Personen, die gemeinhin unter der Bezeichnung ‚Infrastruktur' und ‚Verwaltung' gefasst werden, bleiben häufig unterhalb des Aufmerksamkeitsradars der Wissenschaftler:innen. Doch gerade hier findet sich im Hinblick auf die diskutierten Themen nicht nur wertvolle Expertise, sondern auch ein gewisses Potential, alternative Entwicklungen in den entsprechenden Bereichen voranzutreiben.

Ein Stichwort ist dabei ‚Diamond-Open-Access', welches oft von der wissenschaftlichen Community getragen wird und somit einen nicht-kommerziellen Hintergrund aufweist (vgl. Open Access Network 2025). Die Realisierung des vorliegenden Bandes „Vermessene Philosophie" wurde bewusst nach diesem Modell gestaltet. Sie ist damit konkreter Testfall einer solchen Publikation in den Geisteswissenschaften für das vom Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) geförderte Projekt „Diamond_Thinking", welches als Verbundvorhaben von Mitarbeitenden der Bibliothek des Karlsruher Instituts für Technologie (KIT) sowie der Universitätsbibliothek Stuttgart (Projektleitung) durchgeführt wurde. Ziel dieses Projektes war die Entwicklung von Diamond-Open-Access-(DOA)-Geschäftsmodellen und DOA-Best-Practices für Zeitschriften sowie Bücher und Proceedings. Im Beitrag von Stefan Drößler „Das Projekt Diamond_Thinking" in diesem Band finden sich weitere Informationen dazu und können als Anregung dienen, weitere Veröffentlichungen innerhalb der Philosophie nach diesem Modell zu realisieren.

In analoger Weise enthält das Positionspapier „Elektronisches Publizieren und Bewertung wissenschaftlicher Expertise" (2024) eine Reihe weiterer Vorschläge für Ansatzpunkte zu Veränderungen. Es wurde in diesem Band erneut mit abgedruckt und adressiert vier in diesem Band besprochene Themen, die

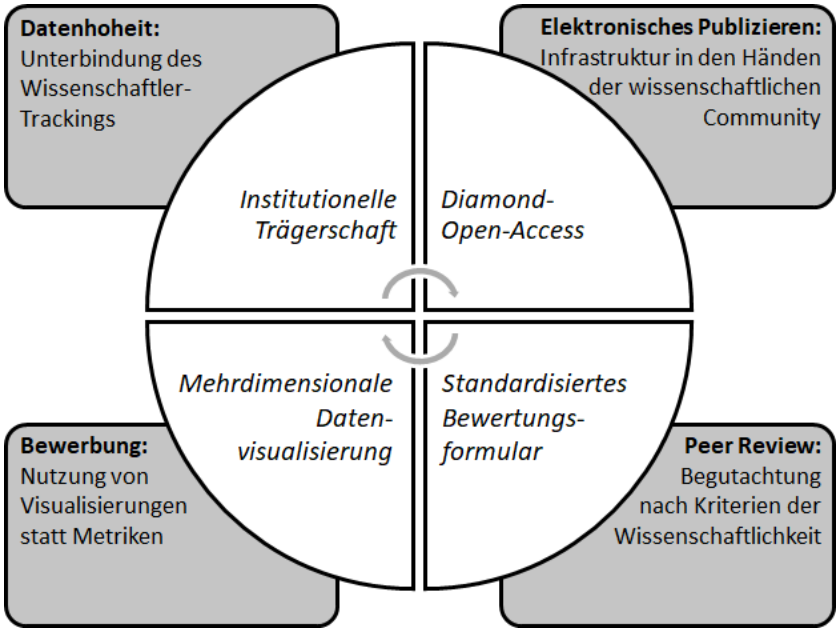


Abbildung 1: Ansatzpunkte für Veränderungen.

eine zentrale Bedeutung für die Art und Weise künftigen wissenschaftlichen Arbeitens haben (vgl. Abb. 1): Wer kann das elektronische Publizieren in den Wissenschaften mitgestalten, wie können sowohl qualitative als auch quantitative Evaluierung verbessert und wie kann die Datensouveränität der Wissenschaftler:innen sichergestellt werden?

Die Vermeidung der angedeuteten Dystopie durch die Änderung der bestehenden Ausgangslage kann allerdings nicht ohne aktives Mitwirken der wissenschaftlichen Community gelingen. Letztlich ist es das Nutzerverhalten und das Evaluationsinteresse der Wissenschaftler:innen, der Philosoph:innen selbst, welche die technologische Forschungslandschaft auf jene Art und Weise gestaltet haben, die nun zu besagten Problemstellungen führen. Halten wir einfach am Altbekannten fest, stärken wir weiterhin die Oligopolstellung der Data-Analytics-Unternehmen. Hier bedarf es letztlich eines Wandels unserer Wissenschaftskultur, um neue Wege einschlagen zu können. Eine solche Veränderung

kann jedoch nicht von außen an die Community herangetragen werden, sondern setzt in entscheidendem Maße auf die gestaltende Rolle ihrer Vertreter:innen selbst.

Auf diese Punkte gehen auch die folgenden Beitragenden des Bandes ein, wenn sie nach Auswegen aus den problematischen Dynamiken einer metrisierten Wissenschaft und den damit verflochtenen Prozessen des elektronischen Publizierens fragen.

Im Beitrag „Perils and Promises: How to Regain Control of Metrics in Academia" diskutiert Hanne Andersen, inwiefern bibliometrische Werkzeuge dazu genutzt werden könnten, sinnvolle Hilfestellungen bei der Eruierung von Werdegängen, Karrierewegen und Profilen von Wissenschaftler:innen zu geben. Andersen beginnt ihre Untersuchung mit einer kritischen Analyse einer Reihe bestehender bibliometrischer Methoden. Sie zeigt auf, wie im Laufe der Entwicklungen zunehmend die Aufgabe der Ermöglichung von Überblickwissen zu wissenschaftlichen Gebieten mit den Fragen nach Bewertung und Auswahl wissenschaftlicher Expert:innen verquickt wurde. In der Folge haben sich diese Werkzeuge tief im Wissenschaftsalltag eingegraben, werden als solche von vielen geschätzt und kontinuierlich genutzt, zugleich aber ebenfalls vehement kritisiert. Andersen analysiert dieses Spannungsverhältnis, indem sie u.a. auf die derzeitigen Hauptschwierigkeiten der Verwendung dieser quantitativen Metriken als Evaluationsinstrumente der Wissenschaft genauer eingeht.

In ihrem Beitrag „Open Science – eine Frage der Wissenschaftskultur" wendet sich Nicola Mößner den Entwicklungen im Bereich der Open Science (OS) zu. Seit einiger Zeit setzen viele Protagonist:innen große Hoffnungen auf OS mit der Erwartung, dass hier valide und praktikable Lösungen erarbeitet werden für die drängenden Probleme, die im Kontext des elektronischen Publizierens und der damit einhergehenden Metrisierung wissenschaftlicher Expertise entstehen. Doch ist das tatsächlich der Fall?

In ihrer kritischen Analyse setzt Mößner mit einer Untersuchung der Frage an, was (derzeit) unter ‚Open Science' (OS) verstanden wird. Mit welchen Herausforderungen und Schwierigkeiten sehen sich Vertreter:innen entsprechender Initiativen

konfrontiert? Deutlich werden diese in den Ausführungen der Wissenschaftstheoretikerin Sabina Leonelli, die eine dezidiert kritische Haltung den gegenwärtigen OS-Vorhaben gegenüber einnimmt. Mößner diskutiert das von Leonelli präsentierte Bild der OS, in welchem ‚Offenheit' v.a. als ein ‚Teilen von Ressourcen' dargestellt wird. Als Alternative dazu postuliert Leonelli epistemische Diversität und Gerechtigkeit als Leitprinzipien für künftige OS-Initiativen. Auf Basis dieser Überlegungen hält Leonelli letztlich eine Neukonzeption der OS für unabdingbar. Dieser Schlussfolgerung widerspricht Mößner in ihrem Beitrag. Die Autorin vertritt demgegenüber die These, dass die entscheidende Stellschraube zur Behebung gegenwärtiger Probleme im Bereich der OS in einer kritischen Auseinandersetzung mit der eigenen (westlichen) Wissenschaftskultur zu suchen ist.

Literatur

Altschaffel, Robert, Michael Beurskens, Jana Dittmann, u.a. (2024) „Datentracking und DEAL – Zu den Verhandlungen 2022/2023 und den Folgen für die wissenschaftlichen Bibliotheken". *RuZ – Recht und Zugang* 5 (1): 23–40. https://doi.org/10.5771/2699-1284-2024-1-23.

DFG-Ausschuss Für Wissenschaftliche Bibliotheken Und Informationssysteme (2021) „Datentracking in der Wissenschaft: Aggregation und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage. Ein Informationspapier des Ausschusses für Wissenschaftliche Bibliotheken und Informationssysteme der Deutschen Forschungsgemeinschaft". *Zenodo*. https://doi.org/10.5281/ZENODO.5900759.

Herb, Ulrich (2018) „Zwangsehen und Bastarde. Wohin steuert Big Data die Wissenschaft?" *Information – Wissenschaft & Praxis* 69 (2–3): 81–88. https://doi.org/10.1515/iwp-2018-0021.

Holzer, Angela (2022) „Die Vermessbarkeit der Wissenschaft: Digitalisierung, wissenschaftliches Publizieren, Verhaltenstracking und Wissenschaftsbewertung". In *Kalibrierung der Wissenschaft: Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis*, hrsg. von Nicola Mößner und Klaus Erlach. Transcript Verlag, S. 163–182. https://doi.org/10.14361/9783839462102-007.

Mößner, Nicola (2024) „Metriken und Expertise". *Zenodo*. https://doi.org/10.5281/ZENODO.14218972.

Mößner, Nicola (2025) „Auf einem Auge blind? ‚Vendor Lock-in' und dessen Folgen in der Philosophie". In *Prae|Faktisch. Ein Philosophieblog*. April 1. https://praefaktisch.de/002e/auf-einem-auge-blind-vendor-lock-in-und-dessen-folgen-in-der-philosophie/.

Mößner, Nicola und Klaus Erlach (2024) „Elektronisches Publizieren und Bewertung wissenschaftlicher Expertise". *Zenodo*. https://doi.org/10.5281/ZENODO.10948471.

Open Access Network: https://open-access.network/informieren/open-access-grundlagen/open-access-gruen-und-gold#c25789, eingesehen am 17. Juli 2025

Retzlaff, Eric (2022) „Wer bewertet mit welchen Interessen wissenschaftliche Publikationen?: Eine Skizzierung des Einflusses kommerzieller Interessen auf die Forschungsoutput-Bewertung". In *Kalibrierung der Wissenschaft: Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis*, hrsg. von Nicola Mößner und Klaus Erlach. Transcript Verlag, S. 139–162. https://doi.org/10.14361/9783839462102-006.

Siems, Renke (2022) „Das Lesen der Anderen. Die Auswirkungen von User Tracking auf Bibliotheken". *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB*, 9(1): 1–25. https://doi.org/10.5282/O-BIB/5797.

CHAPTER 10

Perils and Promises: How to Regain Control of Metrics in Academia

Hanne Andersen

1. Introduction

Academia has become permeated by metrics. Individual scholars highlight the number of their publications to show their productivity, the number of citations of their publications to show their impact, or they highlight their H-index in an effort to describe both their productivity and their impact at the same time. Universities highlight their position in different international rankings in order to attract students as well as researchers. Journals highlight their impact factor in order to attract authors, and, paradoxically, some journals highlight high rejection rates with the same purpose. Committees awarding positions, prizes or funding may draw on, for example, H-indexes or adjusted citation impacts to guide their decisions. Funding bodies draw on metrics aggregated on the institutional level to decide on the level of financial support.

How to cite this book chapter:
Andersen, H. 2026. Perils and Promises: How to Regain Control of Metrics in Academia. In: Mößner, N. and Erlach, K. (eds.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* Pp. 295–317. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.m. License: CC BY 4.0

At the same time as metrics have become increasingly used, they have also become increasingly scorned and frowned upon. Editors protest against classification of journals. Faculty protests against institutional focus on university rankings. International declaration have warned against the use of metrics for assessing individuals, and urged for a return to classical, qualitative expert assessments.

To understand this love-hate relationship that academia has with its metrics, it is necessary first to delve into history. Where do all these metrics come from? Who invented them and why? How did they go haywire? But to move on and develop a more healthy relationship, it is just as important to focus on the present and the future, and to ask what *positive* use metrics may have, and especially how we as academics can regain control of their use.

The chapter is written from the dual perspective of both exploring the history and philosophy of metrics and reflect on whether some of the developed metric-related tools can be used responsibly and sagely. The chapter will therefore start with a brief historical account of the development of various indicators and tools, then take a look at how indicators purport to measure research quality, and finally end with some reflections on alternative uses of the developed scientometric tools.

2. Exploring an endless frontier with limited resources

To understand why metrics have become so popular in academia, it is important to see them in the context of how science has developed as an endeavor. During the 20th century, research became increasingly seen as an important driver of health and prosperity. Countries around the globe increased their allocation of funds for academic research and training of new generations of scholars. However, research has a tendency always to produce more ideas for additional research than what available resources can support. At the mid-20th century, this fundamental tension became vividly

epitomized by the exchange between the US science advisor Vannevar Bush and the US budget director Harold Smith, when Smith suggested that Bush's 1945 report *Science: The Endless Frontier* (Bush 1945); the report that launched big-scale public investments in science as the way to secure peace, health, and prosperity in the future, should instead have been titled *Science: The Endless Expenditure* (Barfield 1997; Godin 2002).

Already in the 1960s, historian of science Derek de Solla Price in his two seminal monographs *Science Since Babylon* (1961) and *Little Science, Big Science* (1963) documented that, at least within the sciences, the number of articles, journals, and graduates was growing exponentially and had been doing so for centuries. Price also emphasized that it was obvious that such an exponential growth could not continue. As he argued, if it did, it would eventually produce "dozens of scientists per man, woman, child and dog of the world population".[1] Thus, on the one hand, the world of research may be characterized by a continuously increasing volume of research activities and research results; on the other hand, resources are limited; both in the form of people to further develop our total body of knowledge through research, and in the form of money to support their research activities.

This creates two distinct challenges. First, faced with growing volumes of research results and growing numbers of scholarly peers, it becomes increasingly challenging to maintain an overview of the scientific record. This challenge therefore presents an increasing need for tools for classification and for searching. Second, faced with a growing need for resources, it becomes more and more challenging to decide on how to distribute the scarce resources that are available. This challenge therefore presents an increasing need for tools for assessing and ranking.

3. The emergence of tools for identifying and ranking

While the tasks of maintaining overviews and of making selections are not identical, they are also not separate, and many of the tools

developed for classifying and identifying scholars etc., or for assessing and ranking have often merged the two, although it varies from tool to tool how it is done, and which consequences it has.

3.1 Science Citation Index/Web of Science

How the approaches for, on the one hand, identifying scholars and their scholarly contributions, on the other hand, for assessing and ranking them, became almost inextricably intertwined is epitomized by the development of the Web of Science (previously the Science Citation Index),[2] that today is one of the most widely used tools for both identifying and ranking scholars, as well as research results.

The index was developed in the 1950s and 1960s by the chemist and information specialist Eugene Garfield. As an information specialist, he had previously worked on indexing, classification, and abstracting and therefore knew that these are highly time-consuming activities, and that the quality of the result depends crucially on the topical expertise of people doing it. He therefore started considering whether it would be possible instead to draw directly on the expertise of the scholarly authors and how they saw their own ideas relate to the ideas of others. Assuming that such relations would be manifested in citations, Garfield imagined that an index based on citations between publications would not only map associations between ideas, it would also provide a much more fine-grained mapping than the categories in traditional indexes (Garfield 1955). For Garfield, in principle such an index could provide a total overview of all existing ideas and their mutual relations, or, as Garfield expressed it "a forerunner of the World Brain" that had been described by H.G. Wells (Garfield 1964), or an index that would offer "the possibility of having all recorded knowledge at one's fingertips" (ibid.). Hence, initially the citation-based index was in this way introduced as an "attempt at universal bibliographical control of science literature" (Garfield 1964) that would make it easy for scholars to track ideas in the

existing literature, and to check how these ideas had been received and criticized (Garfield 1955).

However, from very early on, Garfield also hypothesized that the basic unit of his index, the citation, could be seen not only as a directed relation between publications. Assuming that the number of citations to a given publication indicated how often the ideas in this publication had influenced ideas described in later publications, citations could also be seen as a quantitative indicator of the degree of influence (or impact) that the idea(s) described in a publication had had. This additional idea enabled the aim of the index subtly to drift from tracking ideas through the literature, to evaluating the literature in which the ideas were presented. Further, once the aim of the index had in this way been widened, an obvious next step was to widen it even further to also assess the impact of individual scholars based on the impact of their publications (Garfield 1970).

As Garfield developed and launched the Science Citation Index as a commercial product offered by his company Institute for Scientific Information, he therefore reached out to researchers from the then emerging sociology of science, hoping that they would use the index to study topics such as stratification and reward structures. Although it was soon established that recognition in the form of, for example, prizes or positions at prestigious institutions correlated with citation numbers (Cole and Cole 1967), it was noted from the outset that due to variations at the individual level it was not recommendable to use citation data to inform, for example, promotion and hiring decisions (Charlton 1981; see also Wouters 1999 for a historical exposition, for current warnings see e.g. Hicks et al. 2015; Adler et al. 2009).

The practical implementation of the index created other challenges. Indexing all research publications in the world is obviously a daunting task – if possible at all. First, a distinction needs to be drawn between the scholarly literature to include, and the non-scholarly literature to exclude. Second, the literature to include needs to be identified, across languages and geographical regions. Today, web crawlers may, in principle, be able to find all publications that

have been digitalized and made available on the internet, but before the digital age, the tasks of locating the publications and physically scanning them for references were overwhelming.

To make the task feasible, Garfield therefore made a number of decisions that reflected his interests as director of a commercial company that needed to limit costs, and that had libraries as primary customers. First, arguing that those scholarly journals that are perceived by scholars as especially important typically also receive far more citations than the rest, he argued that limiting the indexing to selected journals would not seriously compromise the coverage of citations substantially. Thus, the idea of an exhaustive coverage of all available scientific literature was abandoned from the outset, and a selection was made that would benefit, for example, broad journals and journals from large fields, at the expense of small journals and journals from small fields that typically carried fewer citations. Further, Garfield initially concentrated on the natural and medical sciences, and not until later were the social sciences and humanities added as separate indexes. Finally, in selecting the journals to include, the decisive factors were not only expert judgment of their perceived importance in their field, but also other factors. For example, an important aspect has remained whether the journal publishes according to schedule. This factor obviously has little to do with the quality of the content. Instead, it is important in order to compute journal impact factors as average citations over fixed timespans.

Thus, while the original ideal had been to provide a complete overview of the scientific record, the aim of the index gradually shifted towards providing measures of the impact of publications, scholars, and journals. Further, the practical implementation of the index was to the advantage of large fields and the disadvantage to small and highly specialized fields.

3.2 ERIH and ERIH PLUS

The Science Citation Index (as well as most similar databases) is focused largely on English-language publications. This has created

special challenges for the humanities and social sciences, especially outside the anglophone world. While during the 20th century English gradually became the *lingua franca* for the natural and health sciences, the humanities and social sciences have remained much more multi-lingual. However, crossing linguistic barriers is often challenging, and research published in less known languages has often been less recognized, especially in the anglophone world.

As a reaction to this visibility problem, the European Science Foundation (ESF) established a European Reference Index for the Humanities (ERIH) in 2008. While the ESF stressed that the aim of the index was to make research of the European humanities published in academic journals in all European languages more visible, they at the same time introduced a categorization of journals that had clear evaluative components. Thus, *Category A* were "high-ranking international journals with high visibility that are regularly cited all over the world and have a very strong reputation among researchers in different countries"; *Category B* were "standard international journals with significant visibility and influence in the various research domains in different countries"; and *Category C* were "high-ranking national or regional journals with recognized scholarly significance" (Fuchs 2014). In some fields, editors protested against this ranking; for example, editors of journals for history of science and technology requested in a joint response to the index that their journals be removed from the index (Andersen et al. 2009).

In 2014, ERIH was transferred to the Norwegian Centre for Research Data that already operated a similar Norwegian register for scientific journals, series, and publishers. This register had been introduced in 2006, also with the dual aim of both making the research production of the country visible, and to incentivize researchers to publish more – and especially to publish in more prestigious publication venues (Komiteene 2010). In this Norwegian system (Frida, later Norsk Vitenskapsindeks, or in popular speak "tellekantsystemet"), as in a similar Danish system ("den danske forskningsindikator") introduced in 2009, all academic journals, series, and publishers are sorted in two ranked categories, with the 20% most influential publishing venues placed in the higher category,

and the rest in the other one (Sivertsen and Schneider 2012). To incentivize publication in the higher category, the system has been used in both countries to distribute funding to the universities based on the number of publications produced in each of the two categories (https://snl.no/tellekantsystemet, accessed 03.11.2024). It has therefore been mandatory for all university-employed researchers to register their publications in a specific database. Each individual publication has then been granted 'publication points' depending on its listing in the higher or the lower category. In both countries, these systems have been introduced by the respective ministries of higher education with the aim of aggregating publication points at the institutional level and to factor these publication points into the computation of the state funding contribution to each individual institution as well as to factor these publication points into the algorithms computing the amount of state support for the institutions (Østerud 2009; Hansen 2023).

The sorting of journals into the two categories was performed by a number of expert groups that each covered a given area. Each journal was assigned to a group, and the experts in the group would then decide whether the journal was categorized into the lower or the higher category. However, this turned out to be far from a simple task. First, it may vary considerably between experts within the same field which journals they consider to be 'better' or 'stronger'. Hence, even within each discipline, some categorizations have been controversial. Second, some journals may cover different fields and be ranked differently by different groups. However, in order to provide one, unique categorization, compromises need to be made for such cross-over journals.[3] Third, highly specialized journals easily get overlooked if their fields are not represented in the relevant committee. Fourth, there is a danger that such a system is 'gamed' by categorizing the journals in which many local researchers publish into the higher-ranking category.[4] Consequently, there has been much debate about the system in Norway and Denmark. In Denmark, it was stopped by the end of 2021 (https://ufm.dk/forskning-og-innovation/statistik-og-analyser/bfi-afviklet, accessed 03.11.2024).

However, the European Reference Index for the Humanities and Social Science still continues (now as ERIH PLUS that includes the social sciences). As of August 2024, the database has indexed more than 10.000 academic journals for which it both provides information about open access status and Plan S compliance, and facilitates an easy search for articles in all journals covered (https://kanalregister.hkdir.no/publiseringskanaler/erihplus/, accessed 03.11.2024). In this sense, ERIH PLUS now serves primarily as a tool for showcasing and identifying literature, not as a tool for assessing its quality, importance, or impact.

3.3 PURE and research information management systems

While the Web of Science and ERIH Plus draw their data on scientific publications, authors, and citations from the publishers' electronic repositories, research information management systems have also emerged and developed that draw or complement their data directly from the individual scholar. The primary feature of these systems is to collect and display information about the achievements of individual researchers on their institutional homepages.[5]

One of the popular systems of this kind is PURE that is provided by the commercial publishing company Elsevier.[6] The database consists both of records imported from databases such as Web of Science, Scopus, and Mendeley, and of records typed in by the individual researcher (or by others doing it on their behalf). Thus, the individual researcher can complement records drawn from the databases with information about, for example, publications in journals or repositories not covered by the databases; manuscripts in preparation; information about prizes, awards, grants, oral presentations, professional service, media coverage etc.; and information about how publications and activities relate to received grants.

In this way, the system enables researchers to improve the visibility of their work. However, as the system has developed, it has

gradually imported more and more of the evaluative features of the databases from which it draws records, such as citations. But since the system is implemented at the institutional level, it varies from institution to institution whether the information about individual researchers is limited to making their achievements visible by listing them, whether it also includes impact indicators like citation numbers, or whether the individual researcher can choose to include such information or not.

4. Assessing "quality" by proxy

The tools that have emerged for identifying and assessing scholarly literature has had immense impact on the assessment in academia. Once upon a time, assessing the originality and overall quality of a scholar's production was seen as requiring careful reading of the produced research by peers in the field. However, as competition in academia has grown more fierce there tends to be more and more applicants for each announced position, and each of these applicants tends to have produced more and more published results. Consequently, it has become increasingly time consuming to make a qualitative assessment, at least in the early stages of, for example, a recruitment process.

At the same time, databases such as the Web of Science or Scopus have enabled increasingly easy access to information about how much scholars have published, and how often their publications have been cited. It has therefore been tempting to alleviate the challenge of the increasing burden of assessments by drawing on publication and citation numbers as indicators of productivity and impact.

Hence, the responsibility for the rise in use of indicators for assessment is not only with the companies that provide these indicators, but also with the scholars themselves who have embraced them as tools for saving time. For example, one of the currently most popular indicators, the H-index, was developed by the physicist J.E. Hirsch as a quantification of scholars cumulative impact that "in a world of limited resources" he found needed for

the purpose of evaluation and comparison, "even if potentially distasteful" (Hirsch 2005).

4.1 The H-index as a mixed indicator

Hirsch introduced the H-index as a mixed indicator that would characterize the scientific output of scholars by taking into account both their productivity and their impact. Hence, the H-index purports to integrate two different phenomena into one unique measure.

Hirsch defined the H-index h of a scholar as the highest number, h, for which the scholar has authored h publications that have each been cited at least h times. Hirsch argued that this number was easy to establish by finding in a database such as Web of Science the publications that a scholar has (co-)authored, and by ordering them according to the number of citations they have each received. From this information a diagram can be drawn with individual papers along the x-axis and number of citations along the y-axis, and papers ordered according to a decreasing number of citations (see below). The dotted line shows where the

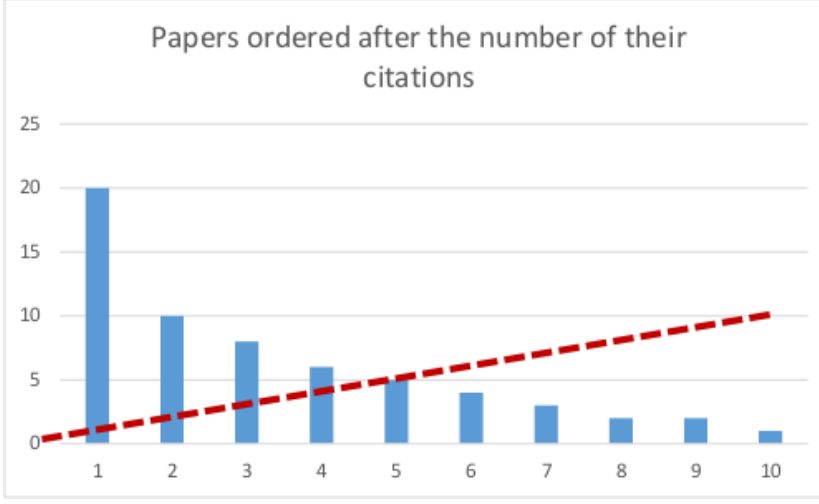


Figure 10.1: Papers ordered after the number of their citations.

number of the papers equals the number of citations. The number of the last paper to reach this line is the H-index.

Due to its simplicity and the relatively easy way in which it could be established, the H-index quickly became immensely popular (Barnes 2017), especially in the natural and health sciences. This became self-reinforcing so that several databases now include the H-index as part of the standard information on individual authors.

However, despite its popularity, the H-index has significant shortcomings. First, a scholar (A) who has published ten papers that have received ten citations each, and another scholar (B) who has also published ten papers that have instead received a hundred or even a thousand citations each, will have the same H-index (10), although most will share the intuition that A seems in some way to be a 'better' or 'stronger' scholar than B. To remedy this shortcoming, alternative indexes, such as the g-index (Egghe 2006) or the e-index, have been proposed that adjust for very highly cited papers (Zhang 2009).

Second, a scholar's H-index never decreases. It does not indicate recent changes in productivity or impacts, if for instance a scholar stops publishing, or stops getting ideas that others cite. Similarly, the H-index naturally increases with career age. To remedy this shortcoming, it has been suggested to use indicators such as the m-quotient that adjusts for career age (Hirsch 2005), or the hIa-index that adjusts for career age and co-authorships (Harzing 2010).[7]

Third, due to the rise in publications with many co-authors, the correlation between the H-index and rewards have declined. To remedy this shortcoming, it has been suggested to introduce a modified H-index based on fractional allocations of citations among authors (Koltun and Hafner 2021).

As argued by the historian and philosopher of science Gingras (2016), it is a typical shortcoming of mixed indicators that they are often insensitive to significant variations in the multiple phenomena that they purport to measure.[8] In case of the H-index, it is insensitive to particularly high-cited papers, as well as to a decline in productivity over time. However, the solution has been

to introduce a plethora of alternative or complementary indicators that are also mixed, and which therefore suffer from their own insensitivities. Thus, despite persistent shortcomings in devising a single indicator that can reliably rank scholars based on both productivity and impact, the "h index zoo" has continued to proliferate. As summarized by Abramo et al. (2013): "Even among actual evaluation practitioners, simplicity and rapidity can prevail over rigor and reliability".

4.2 Normalizing indicators

Another challenge for developing indicators that can assess quality based on publication and citation numbers are the differences in publication and citation practices that exist across disciplines and other contexts.

4.2.1 Productivity

Typically, productivity is assessed by considering how many publications a scholar has (co-)authored in a certain period of time. While it can often be determined quite easily when a scholar started publishing, it remains opaque whether the timespan between the scholar's first and last publication has actually been available for research, or if, for example, parental leaves, sick leaves, periods of unemployment, etc. have reduced the opportunities for conducting research for parts of this timespan. For this reason, comparisons between scholars based on the average number of publications they have produced per year should be handled with caution.

Further, publication practices vary across disciplines, fields, and institutions. Whether the norm is to publish monographs or articles obviously makes a huge difference for the number of publications that a scholar can produce over a given period of time. Similarly, whether it is the norm to publish alone or as a group of collaborating authors can also make a considerable difference for the number of publications that a scholar can (co-)author in a given time. This last point has given rise to intense discussions

on fractional counting of authorships, including the question whether authorship credit should be distributed equally among all authors, or more credit should be granted to e.g. the authors listed first and last on a publications by-line.

4.2.2 Impact

Citation practices also vary widely from field to field. Scholars from scientometrics have therefore introduced field-normalized citation counts to enable cross-disciplinary comparisons. The normalization process consists in restricting comparisons to other publications within the same field, and therefore depends on a classification of all papers into field categories. This means, first, that normalized citation counts are very sensitive to how the databases define their disciplinary categories. Second, it means that normalizations are offered only by databases such as Web of Science or Scopus, not by e.g. Google Scholar.

Web of Science lists for each individual paper its category normalized citation impact computed as the ratio of the paper's citations relative to the expected count for a document of the same type, from the same category, and published in the same year; as well as its percentile in the subject area computed as the percentage of papers of the same type, category, and year that have a lower citation count. For authors, Web of Science offers a beamplot of the category normalized citation impact for all the papers of a particular author which are recorded in the Web of Science database, and it computes the author's overall citation percentile median as the average over all of these papers.

Scopus lists for each individual paper its field-weighted citation impact computed as the ratio of the paper's citations relative to the expected world average for the subject field, publication type, and publication year. For a particular author, the aggregated field-weighted citation impact is averaged over all the papers by this scholar which have been recorded by SciVal.

For both Web of Science and SciVal, the field-weighted or category normalized citation impact averages are highly sensitive

to the author's overall publication strategy. For example, because journal articles tend to receive more citations than book chapters, normalized citation averages tend to reward authors who publish in journals rather than in books. Similarly, since citations from non-English journal publications are often not covered by the databases, normalized citation averages tend to punish authors whose work is directed at non-English-language traditions or audiences.

5. Regaining control

As the use of metrics for assessment purposes in academia has increased, so has the criticism of it. This criticism has come from several different sides, including scientometrics (Hicks et al. 2015; Stephan et al. 2017), statistics (Adler et al. 2009), and history and philosophy of science (Gingras 2016). To support good practices in the use of metrics for assessment purposes, networks of scholars have presented manifestos (Hicks et al. 2015) and declarations (e.g. the Declaration on Research Assessment, DORA https://sfdora.org, accessed 03.11.2024) describing responsible practices. Important points in these guidelines are to stress the relevance of supplementing quantitative assessments with qualitative ones, of drawing on a broad range of different indicators, of avoiding misplaced concreteness and false precisions, and to regularly scrutinize and update the indicators used (see e.g. Andersen 2020; Andersen 2023 for brief overviews).

But as indicators seem to be here to stay, how do we as academics navigate between the increasing use and the intensified criticism of metrics, between the ease of use and the scorn of its simplicity? How can we regain control of the metrics, and develop our practices in ways that make the use of metrics, rankings, and research information management systems less problematic? In this last section of my chapter, I shall briefly describe ways in which scientometric tools can be used for analyzing and describing scholarly profiles and career trajectories.

5.1 Describing profiles and trajectories

While publication numbers, citation numbers, and the mixing of the two into the H-index and its various derivatives may be only poor indicators of the productivity, creativity, and perseverance of scholars, several of the tools provided for establishing these indicators can be quite useful for describing scholarly profiles and career paths.

For example, the diagrammatic representation of scholars' publications ordered in accordance to the number of their citations can be used not only for finding their H-indexes, but also for exploring and comparing their publication strategies and career development in more detail.

To illustrate this point, the diagram below shows two different scholars with the same number of publications, the same number of citations, and the same H index. While scholar B has a "higher peak", that is, B's most influential publications have had a considerably higher impact than any of A's, scholar A has a "fatter tail", that is, A has a much more stable production outcome than B, but with a moderate impact. Which of the two scholars is the "better" one is a question that may not have a simple, unique answer, but will instead crucially depend on the context of evaluation.

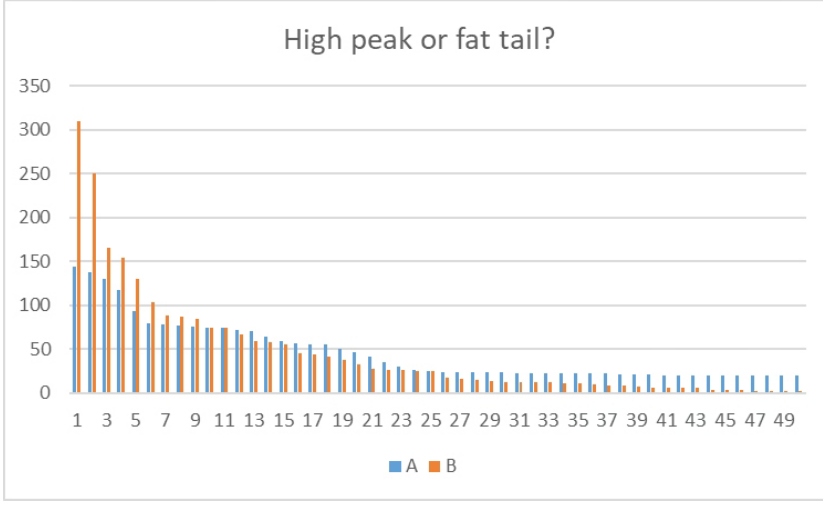


Figure 10.2: High peak or fat tail?

Hence, while the H-index in itself may provide only limited valuable information, a detailed explanation of why the curve has the shape it has may contain a lot of much more interesting information. How can the achievements be described that enabled a higher or a lower peak? For example, is the high peak created by publications that have opened completely new avenues within its field of research, or which have had a substantial uptake across a broad range of fields? Is the lower peak created by publications that have primarily added incrementally to the field's progress, or which have primarily impacted a narrow field with relatively few scholars? Similarly, is the fat tail created by a stable production of incremental contributions to a field's progression, or is it created by participation in a citation cartel whose members meticulously cite each other? Is the thin tail created by a stable production of publications in venues that are well-known for receiving only few citations, such as, for example, book chapters, encyclopedia entries, or journals in non-English languages? Or is it created by a steady production of results that only very few other scholars find interesting at all? Consequently, while the interesting information is not provided by the curves themselves, their shape can prompt and structure valuable reflections on scholars' research and publication strategies, and on how successful they have been in pursuing them.

Similarly, while the average of the normalized (Web of Science) or field-weighted (Scopus) citation impacts of a scholar's publications may itself provide only limited, valuable information, beamplots of the normalized or field-weighted citation impacts for each of the publications enable more interesting reflections. In considering what reasons have led to either very high or very low citation percentiles, the same questions can be used for reflections on and descriptions of research careers, as was described above for the peaks and tails in H-diagrams.

Hence, as a complement to warning and protesting against simplified uses of metrics such as the H-index in the assessment of individual scholars, scholars can indeed embrace the methodologies behind these metrics, albeit for much more nuanced and

insightful descriptions of their careers than what the numbers alone can provide.

5.2 Describing topical scope

Databases such as Web of Science and Scopus not only track citations, they also show how the citations are distributed across fields, disciplines, or academic areas (as these are defined by the database). Typically, this information can be provided for individual papers as well as for the entire production outcome of an author, and it can be visualized as treemaps, piecharts, or doughnut-charts in which each discipline or field is represented by a colored rectangle or share of the piechart. Hence, this can provide an easy overview of whether the impact of a publication or a scholar lies primarily within a particular field or discipline, or if it spreads into many fields or disciplines.

Again, this information may not be valuable in itself, but it may prompt relevant questions for further reflection, e.g. does a specific publication reach its intended audience, does a specific scholar have a profile that matches a highly interdisciplinary or a more disciplinary focused research environment, etc.

6. Final reflections

As described in this chapter, the growth of science during the 20th and 21st century has called for tools for maintaining overviews of a growing scientific record as well as for assessing and ranking scholars competing for scarce resources. Often these two aims have been closely intertwined, and alternating shifts can be noted whether the one or the other aim has been considered decisive for the way in which the tools have been implemented.

Further, the practical implementation of the tools has also often called for pragmatic decisions which have influenced considerations on how well the tools would be suited for different purposes. However, despite widespread criticism of them, especially

when using them for evaluative purposes, their ease of use has made them attractive to many scholars. Unfortunately, what is often overlooked is that while many numerical indicators, such as the H-index or averages over field or category normalized citation counts, may provide information that is too simplified to be of real value, a more careful use of the tools – mindful of their implications – may enable a much more fine-grained overview of scholars profiles, career paths, and range of influence.

In the chapter it has already been argued that the responsibility for the increasing use of indicators lies not only with the companies providing the tools, but also with the scientific community that has embraced theses tools for evaluative purposes in order to save time. Reflecting further along these lines, the scientific community has also to think about how the current use of indicators may influence scholarly activity in the future: If book chapters receive fewer citations than journal articles, will we still have edited collections discussing a specific topic from different perspectives in the future? If encyclopedia entries are read but rarely cited, will such resources be available for students in the future? If not, the worst dangers of metrics may still lie ahead of us.

Notes

[1] Price expected that the growth of science would level off within a few decades. However, neither did he foresee the internationalization of science that would afford much more funding on a global scale, nor did he foresee the explosion in higher education that would make a much larger scientific workforce available. Hence, recent work by Powell, Baker and collaborators (2017). Baker and Powell (2024) has documented that, contrary to Price's prediction, the growth of science has continued and is showing no signs of leveling off.

[2] For the history of the emergence and development of the SCI, see e.g. Wouters (1999).

[3] In Denmark, a rumor told that in the first, internal attempt at merging the categorizations from all 60+ groups, the well-known journals *Science* and *Nature* came out in the lower category only.

[4] There is anecdotal evidence that, if comparing the categorizations of ERIH, the Norwegian "telekant" and the Danish "forskningsindikator", the categorizations of individual journals vary between the three systems, and that the variation often correlates to where scholars from the involved country, or from the categorizing committees, tend to publish.

[5] For example, the page for the author of this chapter can be seen here: https://researchprofiles.ku.dk/da/persons/hanne-andersen, accessed 03.11.2024.

[6] Before being purchased by Elsevier, PURE was developed by Danish universities and university libraries see e.g. Lim 2021.

[7] See e.g. Bornmann et al. (2008) or Harzing (2010). For detailed treatments and comparisons of different indexes see e.g. Andersen (2020), Andersen (2023) for brief overviews.

[8] For a fundamental discussion of heterogeneous indicators and their shortcomings, see e.g. Gingras (2014).

Bibliography

Abramo, G., D'Angelo, C.A. & Viel, F. (2013) "The suitability of h and g indexes for measuring the research performance of institutions". *Scientometrics*, 97: 555–570. https://doi.org/10.1007/s11192-013-1026-4.

Adler, R., Ewing, J. & Taylor, P. (2009) "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)". *Statistical Science*, 24: 1–14. https://www.jstor.org/stable/20697661.

Andersen, H. (2020) "Publish or perish". In: David, M. & Amey, M. (eds.) *SAGE Encyclopedia of Higher Education*. Thousand Oaks: SAGE.

Andersen, H. (2023). "Publish or Perish". *International Encyclopedia of Education*, 4th ed., Elsevier, pp. 158–166. https://doi.org/10.4135/9781529714395.n464.

Andersen, H., Ariew, R., Bag, A.K. & Al., E. (2009) "Editorial: Journals Under Threat: A Joint Response from History of Science, Technology and Medicine Editors". *Centaurus*, 51, 1–4. PMID: 22586774. https://doi.org/10.1017/s0025727300003288.

Baker, D.B. & Powell, J.W. (2024) *Global Mega-Science. Universities, Research Collaborations, and Knowledge Production*. Stanford, CA, Stanford University Press.

Barfield, C.E. (1997) *Science for the 21st century*. Washington DC, AEI Press.

Barnes, C. (2017) "The h-index debate: an introduction for librarians". *Journal of Academic Librarianship*, 43: 487–494. https://doi.org/10.1016/j.acalib.2017.08.013.

Bornmann, L., Mutz, R. & Daniel, H.D. (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*ournal of the American Society for Information Science and Technology*, 59: 830–837. https://doi.org/10.1002/asi.20806.

Bush, V. (1945) *RE: A report to the president by Vannevar Bush, director of the office of scientific research an development*, july 1945.

Charlton, L. (1981) "An index for getting ahead". *New York Times*.

Cole, S. & Cole, J.R. (1967) "Scientific output and recognition: A study in the operation of the reward system in science". *American Sociological Review*, 32(3): 377–390. https://doi.org/10.2307/2091085.

Egghe, L. (2006) "An improvement of the h-index: The g-index". *ISSI newsletter*, 2: 8–9.

Fuchs, M.Ž.J.B.U. (2014) "Bibliometrics: use and abuse in the humanities". In Blockmans, Wim; Engwall, Lars; Weaire, Denis (eds.). *Bibliometrics: Use and Abuse in the Review of Research Performance*, London: Delhi: Portland Press, pp. 107–116.

Garfield, E. (1955) "Citation indexes for science: A new dimension in documentation through association of ideas". *Science*, 122: 108–111. https://doi.org/10.1126/science.122.3159.108.

Garfield, E. (1964) "Science Citation Index"—A New Dimension in Indexing: This unique approach underlies versatile bibliographic systems for communicating and evaluating information. *Science*, 144: 649–654. https://doi.org/10.1126/science.144.3619.649.

Garfield, E. (1970) "Citation indexing for studying science". *Nature*, 227: 669–671. https://doi.org/10.1038/227669a0.

Gingras, Y. (2014) "Criteria for evaluating indicators". In: Cronin, B. & Sugimoto, C.R. (eds.), *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*. Cambridge MA: MIT Press, pp.109–126. https://doi.org/10.7551/mitpress/9445.003.0009.

Gingras, Y. (2016) *Bibliometrics and research evaluation: Uses and abuses*. Cambridge, MA, MIT Press.

Godin, B. (2002) "The number makers: Fifty years of science and technology official statistics". *Minerva*, 40: 375–397. https://doi.org/10.1023/A:1020921620544.

Hansen, T.I. (2023) "Tellekantsystemet. Store Norske Leksikon". https://snl.no/tellekantsystemet. Accessed 2026.

Harzing, A. (2010) *The Publish or Perish book: Your guide to effective and responsible citation analysis*. https://harzing.com/publications/publish-or-perish-book.

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S. & Rafols, I. (2015) "Bibliometrics: the Leiden Manifesto for research metrics". *Nature News*, 520: 429–431. https://doi.org/10.1038/520429a.

Hirsch, J.E. (2005) "An index to quantify an individual's scientific research output". *Proceedings of the National academy of Sciences*, 102: 16569-16572. https://doi.org/10.1073/pnas.0507655102.

Koltun, V. & Hafner, D. (2021) The h-index is no longer an effective correlate of scientific reputation. *PLos One*, 16, e0253397. https://doi.org/10.1371/journal.pone.0253397.

Komiteene, D.N.V. (2010) "Tellekanter [Online]". Available: https://www.forskningsetikk.no/ressurser/fbib/forskning-samfunn/tellekanter/. Accessed 2024.

Lim, M.A. (2021) "Governing higher education: The PURE data system and the management of the bibliometric self". *Higher Education Policy*, 34: 238–253. https://doi.org/10.1057/s41307-018-00130-0.

Powell, J.J., Baker, D.P. & Fernandez, F. (2017) *The Century of Science: The global triumph of the research university*. Bingley, Emerald Group Publishing.

Price, D.J.D.S. (1961) *Science since Babylon*. New Haven, Yale University Press.

Price, D.J.D.S. (1963) *Little Science, Big Science*. New York, Columbia University Press.

Sivertsen, G. & Schneider, J. (2012) "Evaluering av den bibliometriske forskningsindikator". https://api.semanticscholar.org/CorpusID:156930277.

Stephan, P., Veugelers, R. & Wang, J. (2017) "Reviewers are blinkered by bibliometrics". *Nature*, 544: 411–412. https://doi.org/10.1038/544411a.

Wouters, P. (1999) *The Citation Culture*. Amsterdam, Universiteit van Amsterdam.
Zhang, C.-T. (2009) "The e-index, complementing the h-index for excess citations". *PLoS ONE*, 4, e5429. https://doi.org/10.1371/journal.pone.0005429.
Østerud, Ø. (2009) *Hvordan måle vitenskap? Søkelys på bibliometriske metoder*. Oslo, Novus Forlag.

Internet sources

DORA. https://sfdora.org.
https://ufm.dk/forskning-og-innovation/statistik-og-analyser/bfi-afviklet.
https://researchprofiles.ku.dk/da/persons/hanne-andersen.

KAPITEL 11

Open Science – eine Frage der Wissenschaftskultur

Nicola Mößner

Derzeitige in der Praxis des elektronischen Publizierens verankerte Mechanismen sowie das sich zunehmend an quantitativen Maßstäben orientierende Reputationsdenken innerhalb der wissenschaftlichen Community stellen nicht nur die in der Forschung tätigen Wissenschaftler:innen vor wachsende Herausforderungen und Schwierigkeiten, sondern weisen auch das Potential auf, das *System Wissenschaft* insgesamt dysfunktional zu beeinflussen. Auf diese problematische Entwicklung wurde unter anderem von der Deutschen Forschungsgemeinschaft (DFG) aufmerksam gemacht (vgl. Deutsche Forschungsgemeinschaft, AG Publikationswesen 2022).

Die Spitze des Eisbergs der gegenwärtigen Probleme, wenn man so will, bildet dabei das Phänomen des sogenannten *Daten-* oder *Wissenschaftlertrackings* (vgl. DFG-Ausschuss Für Wissenschaftliche Bibliotheken und Informationssysteme 2021;

Zitationsvorschlag für dieses Kapitel:
Mößner, N. 2026. Open Science – eine Frage der Wissenschaftskultur. In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* S. 319–352. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.n.
License: CC BY 4.0

Mößner 2023; 2022a; Siems 2022).[1] Der Zugang über Single-Sign-On-(SSO-)Schnittstellen ermöglicht es den Anbietern von Datenbanklösungen wie *Scopus*, *Web of Science* oder *Dimensions,* die Daten ihrer Nutzer:innen in Profilen zusammenzuführen, die Aufschluss über sämtliche Aktivitäten der Wissenschaftler:innen auf den Plattformen erlauben. Sarah Lamdan (2023) analysiert die Aktivitäten der *Datenkartelle*, wie sie die Betreiber der Plattformen bezeichnet, detailliert. Im Zusammenhang mit dem Problem des Wissenschaftlertrackings erläutert sie: „Clarivate and Elsevier's data analytics products don't just analyze the impact of academic research articles and their authors, they also analyze data collected before and after journals are published. The companies have expanded their analytics to encompass both the prepublication research cycle and the postpublication research evaluation process. […] Their analytics products monetize the entire ‚knowledge production cylce'" (Lamdan 2023: 53). Durch die Registrierung und Zusammenführung jeder Form von Datenspur im Login-Profil – vom Lesen bestimmter Artikel(-abschnitte) über das Publizieren eigener Beiträge (Preprints, fertige Journalartikel, Kommentare) bis zum bloßen Anklicken von Informationsangeboten – werden Wissenschaftler:innen so zu gläsernen Nutzer:innen auf den Plattformen der *Informationsanalysten*, in welche sich die großen Wissenschaftsverlage (Elsevier, Wiley, Springer Nature) zwischenzeitlich umbenannt haben.[2] Das *Profiling* der Anwender:innen (vgl. Mößner und Kitcher 2017), bekannt geworden durch Internetfirmen wie Amazon und Alphabet (Betreiber der Suchmaschine Google), ist längst Alltag im Wissenschaftsbetrieb geworden. In gleicher Weise lässt sich auch in diesem Sektor dieselbe Unwissenheit und / oder Ignoranz der Anwender:innen gegenüber den Auswirkungen dieser Praxis der sekundären Datenverwertung konstatieren (vgl. Zuboff 2019; O'Neil 2016; Morozov 2014; 2015), wie es auch Renke Siems konstatiert: „Wir sehen in den gegenwärtigen Informationsinfrastrukturen einen erheblichen Teil der Prinzipien des Überwachungskapitalismus am Werk, reflektieren dies aber bislang wenig" (Siems 2022: 18).

Von den Nutzer:innen meist unbemerkt, sind durch zunehmende Fusionierungen immer weitreichendere Datenanalyse- und Datenhandelskonzerne entstanden. Im Wissenschaftsbereich decken sie mittlerweile das vollständige Spektrum der Arbeitsschritte (von der Recherche über die Datenauswertung, über das Verfassen eigener Arbeiten und deren Publikation bis hin zur Evaluation wissenschaftlicher Leistungen) ab.[3] Hierdurch wird es einerseits für den individuellen Forscher immer schwieriger, sich dem Zugriff der Datenanalysten zu entziehen. Andererseits werden die im Arbeitskontext gespeicherten und ausgewerteten Datenspuren – ebenfalls von den Nutzenden unbemerkt – mit Informationen aus anderen Lebensbereichen der Anwender:innen verknüpft und als Datenprofile vermarktet. Es geht also längst nicht mehr nur darum, für welche Literatur sich eine individuelle Wissenschaftlerin interessieren mag und was sie wo publiziert. Vielmehr handelt es sich um die Erfassung ihres gesamten informationellen Verhaltens in der digitalen Welt und deren nachfolgende monetäre Verwertung in Form von Profilen für Arbeitgeber, Finanzdienstleister, Versicherungen, staatliche Institutionen und andere mehr. Sarah Lamdan veranschaulicht die Monetarisierung der Nutzerdaten anhand der Konzerne RELX[4] und Thomson Reuters: „RELX and Thomson Reuters call themselves many things – publishers, business services solutions, research platform providers – but at the end of the day they're information dealers. […] RELX and Thomson Reuters don't focus on one information sector. Rather they lead legal information, financial information, academic research, and personal data markets all at once" (Lamdan 2023: 2f.).

Durch dieses umfassende Profiling werden die einzelnen Nutzer:innen zur Verfügungsmasse von Datenverwertern in Unternehmen und in der öffentlichen Verwaltung, sie werden, wie Siems schreibt, „zum Objekt von Entscheidungen anderer" (Siems 2023: 92), die auf der Basis dieser Datenprofile Versicherungspakete schnüren, Finanzdienstleistungen anbieten oder verweigern usw. Die Fragen, ob die dabei zu Grunde gelegten Datenpakete korrekt sind und inwiefern Fehler darin überhaupt und von wem

korrigiert werden können, liegen weit jenseits des Einflussbereichs des datafizierten Subjekts.

Kehren wir mit diesen Hintergrundüberlegungen zurück aus der Dystopie der Datenkartelle. Wie schon angedeutet, stellen diese nur die jüngste Spitze der Probleme dar, welche sich aus der Praxis der digitalisierten epistemischen Arbeitsteilung im Bereich der Kommunikationsprozesse der Wissenschaften ergeben.[5] Lösungen für diese Schwierigkeiten im Kontext des elektronischen Publizierens und der damit einhergehenden Metrisierung von wissenschaftlicher Expertise versprechen schon seit einiger Zeit die Entwicklungen im Bereich der Open Science.

Im nachfolgenden Beitrag soll dieses Lösungsversprechen kritisch in den Blick genommen werden. Zunächst (1) wird erläutert, was unter ‚Open Science' (OS) verstanden wird. Eingegangen wird dabei auf die Ausführungen der UNESCO-Deklaration zu OS aus dem Jahr 2021 sowie auf das Positionspapier der DFG (Deutsche Forschungsgemeinschaft 2022) zu OS, in dem gefordert wird, dass „[d]ie weitere Umsetzung von Open Science […] ausschließlich unter Berücksichtigung wissenschaftskonstituierender Werte gelingen [kann] und […] kein Selbstzweck sein [darf]" (Deutsche Forschungsgemeinschaft 2022: 11). Was heißt es aber, dass von *wissenschaftskonstituierenden Werten* auszugehen ist? Inwiefern unterscheiden sich diese von Werten, die im Kontext der OS-Kultur hochgehalten werden? Diese Fragestellungen leiten über zum nächsten Schritt (2), in welchem Herausforderungen analysiert werden, wie sie sich einerseits aus den derzeitigen technologischen Entwicklungen für OS ergeben sowie andererseits in der wissenschaftstheoretischen Diskussion von Sabina Leonelli (2023; 2022; Leonelli u.a. 2021; Staunton u.a. 2021; Levin u.a. 2016) prominent hervorgehoben werden. Leonelli kritisiert die OS-Bewegung für ihr Verständnis von Offenheit und den Fokus auf das Teilen von Ressourcen. Sie glaubt, dass letzteres weder eine notwendige noch eine hinreichende Bedingung für eine zuverlässige und verantwortungsvolle OS-Strategie darstelle (Leonelli 2023: 8). Demgegenüber postuliert sie epistemische Diversität und Gerechtigkeit als Leitprinzipien für künftige OS-Initiativen.

Im nachfolgenden Beitrag soll analysiert werden, inwiefern diese von Leonelli vorgeschlagenen Werte den Anforderungen entsprechen, die Forschungsförderer wie die DFG für die Realisierung von OS-Vorhaben formulieren. Dieser kritische Abgleich wird uns zur Frage nach der gegenwärtigen Wissenschaftskultur (3) führen, die derzeit einer vollen Potentialentfaltung von OS-Vorhaben hinderlich im Weg zu stehen scheint. Im vorliegenden Beitrag wird die These vertreten, dass gerade hier die wesentliche Stellschraube für gelingende OS-Initiativen zu finden ist, welche das Potential haben, dysfunktionalen Entwicklungen wie dem Datentracking der Informationsanalysten nachhaltig entgegenzuwirken.

1. Was ist ‚Open Science'?

Beginnen wir mit einer Begriffsbestimmung: was wird unter ‚Open Science' (OS) verstanden? Eine Definition findet sich in der Empfehlung der UNESCO zu OS. Hier wird OS bestimmt als „[…] an inclusive construct that combines various movements and practices aiming to make multilingual scientific knowledge openly available, accessible and reusable for everyone, to increase scientific collaborations and sharing of information for the benefits of science and society, and to open the processes of scientific knowledge creation, evaluation and communication to societal actors beyond the traditional scientific community. It comprises all scientific disciplines and aspects of scholarly practices, including basic and applied sciences, natural and social sciences and the humanities, and it builds on the following key pillars: open scientific knowledge, open science infrastructures, science communication, open engagement of societal actors and open dialogue with other knowledge systems" (UNESCO 2021: 7).

Im Einzelnen betrachtet, handelt es sich um ein Bündel von Technologien und Maßnahmen, mit deren Hilfe laut der Definition drei Zielsetzungen realisiert werden sollen:

(a) Wissenschaftliche Erkenntnis soll besser und leichter verfügsowie nachnutzbar gemacht werden, daher soll sie für alle offen

zugänglich sein. Dieser Punkt wurde weitergehend präzisiert, indem klargestellt wurde, dass völlige Offenheit nicht immer zielführend, manchmal sogar schädlich sein könne. Entsprechend wird davon gesprochen, dass wissenschaftliche Erkenntnis lediglich „so offen wie möglich" sein solle: „Access to scientific knowledge should be as open as possible, but sometimes access may need to be restricted, for example to protect human rights, confidentiality, intellectual property rights, personal information, threatened or endangered species, and sacred and secret indigenous knowledge. Open science encourages scientists to develop tools and methods for managing data so that as much data as possible can be shared, as appropriate" (UNESCO und Canadian National Commission for UNESCO 2022: 5).

(b) Die wissenschaftliche Zusammenarbeit und das Teilen von Informationen sollen gefördert werden zum Wohle der Wissenschaft und der Gesellschaft.

(c) Schließlich sollen wissenschaftliche Prozesse der Erkenntnisgewinnung, -bewertung und -kommunikation geöffnet werden für gesellschaftliche Akteure jenseits der traditionellen wissenschaftlichen Gemeinschaft.

OS bezieht sich auf alle wissenschaftlichen Disziplinen gleichermaßen. Grundlagenforschung ist genauso angesprochen wie angewandte Wissenschaften. Darüber hinaus wurden von der UNESCO vier Säulen formuliert, auf denen OS ruht:

(i) Offene wissenschaftliche Erkenntnis bildet die erste Säule mit der schon erwähnten Einschränkung, dass damit „so offen wie möglich, so geschlossen wie nötig" gemeint ist, denn nicht immer ist es hilfreich, alle Informationen in die Öffentlichkeit zu tragen. Zur Illustration dieses Punktes denke man an Patientenakten im Rahmen medizinischer Forschung.

(ii) Gefordert wird darüber hinaus eine offene wissenschaftliche Infrastruktur. Das klassische Beispiel hier wären Repositorien an Hochschulbibliotheken, über welche Schriften der Wissenschaftler:innen zugänglich gemacht werden. Daneben ist unter diesem Punkt aber auch so etwas gemeint wie z.B. die Vergabe von sogenannten *digital object identifiers* (DOIs), mit deren Hilfe Dateien im Internet langfristig auffindbar gemacht werden.

Ziel ist letztlich der langfristige und nachhaltig unbeschränkte Zugang zu wissenschaftlichen Informationen. Entscheidend hierfür ist, dass nicht-kommerzielle Lösungen implementiert werden, die verhindern, dass Marktmechanismen den Zugang zu Informationen behindern oder unterbrechen. Man denke exemplarisch an die Übernahme des Kurznachrichtendienstes *Twitter* und die nachfolgenden Restriktionen im Online-Zugriff auf bestehende Inhalte durch den neuen Inhaber Elon Musk.

Relevant ist ferner die Definition und Einhaltung von Standards z.B. hinsichtlich der Vergabe von Metadaten für Dateien. Angestrebt wird dabei die Vermeidung von Problemen, wie sie u.a. durch inkompatible Software u.Ä. entstehen können, ein Beispiel für derartige Schwierigkeiten wären in *Flash* programmierte Internetseiten, die in neueren Webbrowsern nicht mehr korrekt angezeigt werden können. Sollen Daten langfristig offen zugänglich bleiben, müssen solche Fehlentwicklungen verhindert werden.

(iii) Des Weiteren bedarf es einer Offenheit für gesellschaftliche Akteure und einen offenen Dialog mit Vertreter:innen anderer Wissenssysteme. Hier wird eine Vernetzung und Zugänglichmachung von Informationen über die wissenschaftliche Community hinaus angestrebt. Ein klassisches Beispiel für Vorhaben dieser Art sind die sogenannten *citizen sciences*, bei denen Laien auf unterschiedliche Weise in den Forschungsprozess mit eingebunden werden.[6] Ein typisches Projekt in diesem Rahmen wäre *galaxy zoo*[7], ein Vorhaben, bei dem wissenschaftliche Laien helfen, Fotografien von Galaxien – gewonnen aus der *Sloan Digital Sky Survey* – nach deren Formen (z.B. Balken- vs. Spiral-Galaxien) zu klassifizieren.

(iv) Schließlich geht es darum, die Diversität in der wissenschaftlichen Praxis zu erhöhen, indem Vertreter:innen indigener, marginalisierter, lokaler Gruppen mit ihrem Wissen, ihren Vorstellungen, Fragen und Bedarfen in die wissenschaftliche Kommunikation mit eingebunden werden. Auf die Relevanz solcher integrativen Kommunikationsansätze haben beispielsweise feministische Wissenschaftstheoretiker:innen hingewiesen, wie Helen Longino (2013) ausführt. Erst auf diese Weise konnte z.B. verdeutlicht werden, dass es unplausibel ist, die Wirksamkeit von

> Medikamenten allein am männlichen Stereotyp zu untersuchen, obwohl sich Krankheitsverläufe im weiblichen Körper oft anders zeigen und ggf. einer anderen Therapieform bedürfen.

OS bedeutet also sowohl eine Öffnung für Akteur:innen innerhalb der wissenschaftlichen Gemeinschaft als auch eine stärkere Vernetzung mit wissenschaftsexternen Gruppierungen, womit grundsätzlich der Bestrebung einer stärkeren Demokratisierung der Wissenschaft nachgekommen werden soll.

Dies spiegelt sich auch in den Kernwerten wider, die den Geist der OS im Sinne der UNESCO tragen: Verwiesen wird dabei auf (1) Qualität und Integrität, (2) Allgemeinnutzen, (3) Gleichheit und Fairness, (4) Diversität und Inklusivität (UNESCO 2021: 17). Orientiert an diesen Werten sollen die Zielsetzungen von OS verwirklicht werden, wobei die UNESCO-Empfehlung sich als ein Rahmenwerk versteht, innerhalb dessen entsprechende Projekte initiiert und Schritt für Schritt verfolgt werden können. Die Mitgliedsstaaten werden aufgefordert, sowohl in die Entwicklung der für OS notwendigen technischen Infrastruktur zu investieren als auch insbesondere für einen Kenntnistransfer zu OS und die Ausbildung einer entsprechenden Wissenschaftskultur Sorge zu tragen.

Deutlich wird, dass OS ein multidimensionales Vorhaben innerhalb der Wissenschaftslandschaft darstellt, das auf eine langfristige und nachhaltige Öffnung des Wissenschaftsbetriebs abzielt. Realistisch betrachtet, muss jedem Beteiligten klar sein, dass diese Umgestaltung des bisher vor allem auf Wettbewerb ausgerichteten und an Marktmechanismen orientierten Bereichs epistemischer Praxen in unseren Gesellschaften kein Ad-hoc-Vorhaben sein kann, sondern über längere Zeitspannen vorangetrieben werden muss. Auch ist die Annahme plausible, dass die Entwicklung neuer Technologien schneller erfolgen wird als die Umgestaltung der zugehörigen Wissenschaftskultur, d.h. der Verhaltens- und Denkweisen beteiligter Wissenschaftler:innen. Dass es hier also zu Verzögerungseffekten kommen wird, bedingt durch das Festhalten an tradierten Paradigmen wissenschaftlicher Forschungsarbeit, steht zu erwarten.

Trotz dieser naheliegenden Einsichten in die Entwicklungsdynamiken von OS-Initiativen muss festgestellt werden, dass die metawissenschaftliche Reflexion gerade hier mit ihren kritischen Analysen ansetzt. Zum einen ist dies natürlich wünschenswert, um auf Schwierigkeiten aufmerksam zu machen, die bei der Realisierung von OS-Vorhaben auftreten und behoben werden müssen. Zum anderen ziehen einige metawissenschaftliche Untersuchungen aber auch Konsequenzen nach sich, die einer weiteren Entfaltung von OS-Vorhaben eher hinderlich zu werden scheinen. Im Folgenden soll dies exemplarisch an wissenschaftstheoretischen Betrachtungen zum Aspekt der Offenheit von OS-Entwicklungen aufgezeigt werden.

2. Herausforderungen für Open Science

In den obigen Ausführungen ist bereits darauf hingewiesen worden, dass eine Schwierigkeit für OS-Initiativen in der genauen Gewichtung des Grads an Offenheit besteht, der im Zusammenhang mit der Realisierung entsprechender Projekte gefordert werden soll. Die Devise „So offen wie möglich, so geschlossen wie nötig" lässt zweifelsohne eine hohe Bandbreite an Interpretationen zu.

Zusammenhängend damit kann gefragt werden, was unter ‚Offenheit' überhaupt verstanden werden soll. Nadine Levin et al. (2016) sind dieser Schwierigkeit in ihrer Analyse nachgegangen. Anhand von exemplarischen Interviews mit Wissenschaftler:innen aus dem Bereich der Biomedizin haben sie versucht zu eruieren, was im Wissenschaftsalltag von den Akteur:innen unter dem Aspekt der Offenheit verstanden wird und welche Herausforderungen für OS-Initiativen von den Forschenden gesehen werden. Als Grundverständnis für OS setzen die Autor:innen dabei folgende Definition an: „Here, we take the Open Science movement to broadly include the access to, dissemination of, and reuse of publications, data, materials, and methods" (Levin u.a. 2016: 128). In den Interviews wurden dann insgesamt sieben Kernthemen erfasst, welche das Verständnis, die Erfahrung und

die Praxen der Offenheit im Arbeitsalltag der befragten Wissenschaftler:innen formen (ebd.: 132). Hierzu zählen Aspekte der folgenden Art: (1) die zeitnahe Bereitstellung und Zugänglichmachung von Forschungskomponenten[8], (2) Standards für das Format und die Qualität derselben, (3) Metadaten und Annotationen, (4) Zusammenarbeit mit Kolleg:innen und Communities, (5) freie Wahl der Wege und Strategien für die Verbreitung von Forschungsergebnissen, (6) transparente Peer-Review-Systeme, (7) Zugang zu Forschungskomponenten in nicht-westlichen und / oder nicht-akademischen Kontexten (ebd.: Tab. 2, 137). Hier zeigt sich, dass von den Wissenschaftler:innen keine Definition des Begriffs geliefert wurde, sondern vielmehr sowohl Anforderungen und Wünsche an OS-Initiativen genannt wurden. Diese lassen Rückschlüsse darauf zu, wo Probleme und Schwierigkeiten bei der Realisierung solcher Vorhaben entweder selbst erfahren oder zumindest antizipiert wurden – z.B. in der Feststellung, dass ohne die Vergabe sinnvoller Metadaten das Speichern und Wiederauffinden von Dokumenten in großen Datenmengen etc. kaum möglich ist.

Resümierend halten Levin et al. in ihrer Studie fest, dass ‚Offenheit' im Arbeitsalltag der Forschenden an unterschiedlichen Punkten verschiedene Bedeutungen annehmen könne: „What our findings do highlight is how, in order to understand the implications of Open Science policies, close attention must be paid to the variety of forms that openness can take in different stages and locale of research practice" (Levin u.a. 2016: 137).

Zu einem ähnlichen Ergebnis hinsichtlich der Diversität der Dimensionen von Offenheit kommt Kevin C. Elliott (2022). Er schlägt zur besseren Handhabung des Themas eine Taxonomie der Transparenz oder Offenheit vor. Unterschieden werden von ihm die Dimensionen des Zwecks (warum Offenheit?), der Zielgruppen (offen für wen?), des Gehalts (was soll offen verfügbar gemacht werden?) und der Methodik (wie soll Offenheit erreicht werden?). Bezüglich der Methodik kommen dann Fragen nach dem Zeitrahmen, den Akteur:innen, den Mechanismen und den Wegen der Öffentlichmachung ins Spiel (ebd.: Abb. 1). Angesprochen werden

darüber hinaus auch Gefahren, die mit einer Erhöhung von Offenheit bzw. Transparenz einhergehen können. Hier kommen wir auf den Zusatz „so geschlossen wie nötig" zurück.

Zur *Komplexität* des Offenheitsaspekts tritt dessen *graduelle Ausformung*. Beide Punkte, die sich in der metawissenschaftlichen Reflexion deutlich abzeichnen, sollen nachfolgend weitergehend thematisiert werden.

2.1 Gradualität: Offenheit vs. Transparenz

Dass ein Mehr an Offenheit als erstrebenswert angesehen wird, hängt grundsätzlich damit zusammen, dass *Offenheit* selbst als ein Wert innerhalb der wissenschaftlichen Gemeinschaft betrachtet wird. Petra Gehring[9] verwies im Rahmen der Podiumsdiskussion „Re-Kalibrierung – Wissenschaftskommunikation im digitalen Wandel"[10] auf diese positive Konnotation des Begriffs: *wer will nicht gerne für Offenheit einstehen,* ließe sich dementsprechend zugespitzt fragen.

Gleiches kann auch für den Begriff der Transparenz festgehalten werden, über welchen Gehring im Zuge einer kritischen Auseinandersetzung mit einer Ethik der Digitalität Folgendes schreibt: „Das Wort hat einen sympathischen Klang. Es erinnert an das angelsächsische demokratietheoretische Leitbild einer permanenten Kontrolle (oder wenigstens Kontrollierbarkeit) des Regierungs- und administrativen Handelns durch alle Bürgerinnen und Bürger, gleichsam von unten"(Gehring 2023: 633f.).

Diese positive Wertsetzung von Offenheit und Transparenz mache es Individuen und / oder Gruppen schwer, sich der Forderung nach Offenlegung bzw. Zugänglichmachung von Daten etc. zu entziehen.[11] Warum aber sollte man das auch tun wollen? Sprich, warum kann es sinnvoll sein, Transparenz und Offenheit kritisch zu sehen und also zumindest keinen vollständigen Grad an Transparenz bzw. Offenheit anzustreben?

Elliott (2022) führt im Zusammenhang mit der von ihm erarbeiteten Taxonomie folgende mögliche Gefahren von Transparenz

bzw. Offenheit auf: „wasting scarce resources, slowing down science, harming companies, violating privacy, generating inappropriate skepticism, creating a false sense of trust, causing confusion, facilitating efforts to harass or mislead" (ebd.: Tab. 1). Greifen wir einmal den Punkt „causing confusion" heraus. Hier geht es um die triviale Einsicht, dass mehr Offenheit auch zu einem Mehr an Informationen führt. Mehr Informationen können aber von verschiedenen Akteur:innen ggf. nicht richtig beurteilt bzw. in ihrer Relevanz hierarchisiert werden, sodass sich der Einzelne letztlich in der Informationsflut zu verlieren droht.[12] Dieser Punkt ließe sich fortführen, indem man feststellt, dass *mehr Offenheit* durchaus nicht zu *mehr Transparenz* führen muss. Diese beiden Konzepte stets in eins zu setzen, wie es exemplarisch in dem Text von Elliott (2022) erfolgt, wird somit fraglich.

Dass es ferner einer kritischen Evaluation der zu veröffentlichenden Daten bedarf, ist dagegen den meisten schnell einsichtig. Je nach Forschungsgegenstand und akademischer Disziplin können sensible Daten in der einen oder anderen Form essentieller Teil der eigenen Arbeit sein, z.B. in Form von Patientendaten im Bereich der medizinischen Forschung. Dass diese nicht einfach mit der gesamten Öffentlichkeit geteilt werden dürfen, nur weil man der Devise nach einem Mehr an Transparenz und Offenheit folgen will, versteht sich eigentlich von selbst.

Die DFG (Deutsche Forschungsgemeinschaft 2022) hält in ihrem Positionspapier zu OS fest, dass eine vollständige Offenheit auch deshalb nicht immer zielführend sei, weil „[e]s [...] wissenschaftsdienlich sein [kann], dass z. B. ein geschützter Raum für wissenschaftliche Diskurse oder hinsichtlich der Begutachtung und Bewertung von Projekten gewährleistet ist. Weiterhin können z. B. auch (vertrags-)rechtliche Erwägungen und Patente eine einschränkende Rolle in diesem Kontext spielen" (ebd.: 5). Wissenschaft ist nach wie vor ein hierarchisches System, geprägt von vielfältigen Abhängigkeiten. Völlige Transparenz mag dazu führen, dass Kritik im Kontext von Gutachten etc. nicht frei geäußert werden würde, weil nachteilige Auswirkungen für die eigene berufliche Laufbahn befürchtet werden.

Ein anderer Punkt, der im Zusammenhang mit der Frage nach dem Grad der Offenheit wissenschaftlicher Ergebnisse im digitalen Raum zunehmend eine Rolle spielt, tritt in seiner vollen Tragweite dagegen erst langsam zu Tage. In seinem Buch „The 4th Revolution" (2014) thematisiert Luciano Floridi die Fragen nach der Entwicklung und dem Status künstlicher Intelligenz (KI). Er konstatiert, dass eine sogenannte *starke* KI[13] gegenwärtig weder bestehe noch deren Realisierung in naher Zukunft zu erwarten sei. *Schwache* Formen von KI seien dagegen bereits ubiquitär und ihre Nutzer:innen sorgten kontinuierlich dafür, dass diese Systeme immer enger miteinander verzahnt und dadurch erfolgreicher arbeiten würden. Floridi bezeichnet diese humanoide Anpassung an Informations- und Kommunikationstechnologien (IuK-Technologien) als „enveloping": „[…] we were actually changing the world to fit light, reproductive AI. ICTs are not becoming more intelligent while making us more stupid. Instead, the world is becoming an infosphere increasingly well adapted to ICT's limited capacities" (Floridi 2014: 143). Es lässt sich mutmaßen, dass sich dieser Vorgang des *Envelopings* nun auch im Zusammenhang mit OS-Initiativen wiederfindet bzw. bestimmte Entwicklungen in diesem Bereich so beschrieben werden können.[14]

Seit Kurzem machen KI-Anwendungen, die nach dem Prinzip der Large Language Models (LLM) à la ChatGPT funktionieren, mehr und mehr von sich reden. Entscheidend für den wachsenden Erfolg dieser Technologien ist es, dass sie ein umfangreiches Trainingsprogramm absolvieren können. Die Algorithmen der LLMs errechnen auf statistischer Basis die Wahrscheinlichkeit von Wortzusammenstellungen, sprich, mit welcher Wahrscheinlichkeit welche Wortkombinationen in der jeweiligen Sprache auftreten. Abfragen (sogenannte *Prompts*) an diese Systeme werden auf der Basis solcher statistischen Auswertungen generiert. Aufgaben, wie Texte in einem bestimmten Stil zu schreiben oder Antworten auf spezielle Fragen zu geben, können diese Systeme umso besser bewältigen, je mehr Beispielfälle sie kennengelernt haben. Und hier kommen nun die OS-Vorhaben ins Spiel: die Forderung nach mehr Offenheit und Transparenz

bedeutet im Umkehrschluss auch, dass mehr Informationen z.B. via Open-Access-Initiativen im Netz verfügbar gemacht werden – und folglich, dass mehr Datenmaterial zum Training der KI-Anwendungen zur Verfügung steht. Letztere folgen aber selbst nicht den Grundsätzen von OS, sondern stellen zumeist Produkte kommerziell arbeitender Unternehmen dar. Zugang und Nutzungsmöglichkeiten sind dementsprechend auf Käufer:innen beschränkt.[15]

Neben diesen Ausschlussmechanismen und, damit einhergehend, mit einer sich weiter öffnenden digitalen Kluft zwischen den Anwender:innen implizieren diese Technologien auch eine Reihe von Herausforderungen und teilweise schwerwiegenden Problemen für den Wissenschaftsbetrieb selbst. Erstere betreffen insbesondere ihre Nutzung im Zusammenhang mit der Erbringung von Prüfungsleistungen. In einer aktuellen Studie des *Bayerischen Forschungsinstituts für Digitale Transformation (bidt)* zur Nutzung von generativen KI-Systemen durch Schüler:innen und Studierende halten die Autor:innen u.a. fest; dass „[e]ine [...] Herausforderung [...] in einer adäquaten Leistungsbewertung von Lernenden in Zeiten generativer KI [bestehe], denn es ist oft schwierig und in vielen Fällen unmöglich festzustellen, ob Leistungen von Lernenden selbst oder von einer generativen KI erbracht wurden. Darum sollte dem Wunsch vieler Befragten Rechnung getragen werden und Prüfungsformate an Schulen und Hochschulen angepasst werden" (Schlude u.a. 2024). Ein weiteres Ergebnis dieser Studie, das Anlass zur Sorge bereitet, besteht darin, dass jeweils nur etwa die Hälfte der befragten Schüler:innen und Studierenden angaben, sich bewusst zu sein, dass die von der KI gelieferten Ergebnisse „faktisch falsch" sein können (50% der Schüler:innen, 56% der Studierenden). In vergleichbarem Maße sind sich die Befragten darüber im Klaren, dass die gelieferten Ergebnisse „unvollständig, unausgewogen, widersprüchlich, und/oder unangebracht" sein könnten (40% der Schüler:innen, 58% der Studierenden). Das bedeutet im Umkehrschluss wiederum, dass sich jeweils ungefähr die Hälfte der befragten Schüler:innen und Studierenden dieser Probleme eben nicht bewusst ist.[16]

Schwerwiegende Probleme entstehen, wenn generative KI-Systeme zur Erstellung sogenannter *Fake Publikationen* eingesetzt werden (vgl. Vallbracht 2024). Statt der Veröffentlichung von gründlich recherchierten und geprüften Daten würden Wissenschaftler:innen zunehmend der Versuchung erliegen, schnell neue Texte mit den genannten Technologien zu erzeugen, um dem Druck der allgegenwärtigen *Publish-or-Perish*-Kultur gerecht werden zu können. Es entsteht somit das Problem, dass generative KI-Systeme einerseits selbst zur Verbreitung von fehlerbehaftetem und / oder schlicht falschem Forschungsoutput beitragen. Andererseits werden diese Systeme durch die Öffentlichmachung solcher Falschinformationen wiederum selbst mit falschen Daten trainiert und iterieren die Unwahrheiten weiter.

Kommt man also der Forderung der OS-Initiativen nach einer gesteigerten Offenheit von wissenschaftlichen Daten, Texten u.a. unkritisch nach, wird man diese skizzierten Entwicklungen notgedrungen weiter mit vorantreiben, was weder im Sinne des Wissenschaftssystems noch im Sinne des einzelnen Forschenden sein kann.

Schließen wir diese Sammlung zu Problemen, die sich aus dem graduellen Aspekt der Offenheitsforderung der OS-Initiativen ergeben, durch den Verweis auf die eingangs bereits angesprochene Schwierigkeit des Daten- bzw. Wissenschaftlertrackings. Auch diese für System und Individuum nachteilige Entwicklung kann durch ein unreflektiertes und (zu) weitreichendes Öffnen von Datenquellen zunehmend verschärft werden. Auf diesen Umstand weist wiederum die DFG (Deutsche Forschungsgemeinschaft 2022) hin, wenn festgehalten wird, „dass die völlige und unregulierte Transparenz aller Prozesse und Daten ungewollt auch unerwünschte Kontrollmechanismen ermöglichen kann, sei es durch kommerzielle, staatliche oder supranationale Akteure“ (ebd.: 6).

2.2 Komplexität: OS-Ziele vs. wissenschaftliche Praxis

Neben der Frage nach dem Grad wünschenswerter Offenheit ergeben sich Herausforderungen für OS-Initiativen auch im Hinblick

auf die Komplexität der in den Zielsetzungen formulierten Leitideen. In diesem Zusammenhang sei auf die Analysen von Sabina Leonelli zum Thema hingewiesen.[17]

Leonelli hat sich u.a. kritisch mit der Zielsetzung der Erhöhung von Diversität innerhalb der Wissenschaftslandschaft durch OS-Initiativen auseinandergesetzt. Wie erläutert, gehört es zu den angestrebten Entwicklungen, Vertreter:innen lokaler, indigener, marginalisierter Gruppen stärker in Forschungsaktivitäten einzubinden. Hier geht es gleichermaßen um den Einbezug von Kenntnissen als auch von Bedarfen dieser Bevölkerungsgruppen.

Darüber hinaus sollen mehr für wissenschaftliche Forschung notwendige Ressourcen (z.B. Laborergebnisse etc.) auch Akteur:innen in Ländern zugänglich gemacht werden, die nicht über die finanziellen Möglichkeiten verfügen, Datenzugänge über klassisch kommerzielle Anbieter zu subskribieren.

Es geht also gleichermaßen um die Öffnung des Wissenschaftssystems für wissenschaftliche Laien[18] als auch um die Einbindung von Wissenschaftler:innen aus unterrepräsentierten (Dritte-Welt-)Staaten. In der Präambel der UNESCO-Empfehlung zu OS heißt es dementsprechend: „Acknowledging that open science should not only foster enhanced sharing of scientific knowledge solely among scientific communities but also promote inclusion and exchange of scholarly knowledge from traditionally underrepresented or excluded groups (such as women, minorities, indigenous scholars, scholars from less-advantaged countries and low-resource languages) and contribute to reducing inequalities in access to scientific development, infrastructures and capabilities among different countries and regions.

Also recognizing that open science respects the diversity of cultures and knowledge systems around the world as foundations for sustainable development, fostering open dialogue with indigenous peoples and local communities and respect for diverse knowledge holders for contemporary problem solving and emergent strategies towards transformative change“ (UNESCO 2021: 5).

Leonelli wendet sich diesen Ansprüchen der OS-Bewegung kritisch zu und hinterfragt, ob diese weitreichenden Zielsetzungen

mit den gewählten Methoden der OS-Initiativen realisiert werden können (vgl. Leonelli 2022). Dabei zeichnet sie nach, dass sich als Grundidee der Bewegung v.a. das Teilen von Ressourcen manifestiert habe. Sie hält fest, dass „[…] many of them [OS-Stakeholder, NM] agree on the importance of a seamless flow of research materials, resources and information […]. Openness thus construed revolves around the idea of unlimited access: the desire to make any research element available to anyone who may wish to use it as soon as possible after its creation" (Leonelli 2023: 17). Gerade diesen Fokus auf dem Aspekt des Teilens unterschiedlicher Ressourcen sieht Leonelli jedoch als eher kontraproduktiv an. In vier Fallstudien zeigt sie auf, inwiefern der Wunsch nach mehr epistemischer Diversität mit der Methode des Teilens von Ressourcen in Konflikt gerät.

Die Autorin verweist auf empirische Studien, die u.a. zeigen, dass die OS-Bewegung bei Mitgliedern der genannten sozialen Gruppen und bei Wissenschaftler:innen entsprechender Staaten gerade wegen dieses Fokus' einen schlechten Stand habe und am Ziel, diese Personen stärker in Forschungsprozesse einzubinden, vorbeilaufe. Beispielsweise würden Initiativen zu Open Data deshalb ins Leere laufen, weil potentielle Datenlieferanten aus der Dritten Welt eine berechtigte Sorge vor Benachteiligungen hätten (Leonelli 2023: 20ff.). Zwar könnten sie wertvolle Daten – z.B. im Zusammenhang mit der Ausbreitung des Corona-Virus' sowie von mit diesem zusammenhängenden Krankheitsverläufen, Virusmutationen etc. – bereitstellen, doch bestünde die berechtigte Sorge, dass diese Zulieferung von forschungsrelevanten Daten ohne eine entsprechende Anerkennung unter Wissenschaftler:innen der Ersten Welt bleibe. Im Kontext der Corona-Pandemie wurde diesem Umstand Rechnung getragen, indem eine für den Austausch relevanter Daten genutzte Datenbank ihre Nutzer:innen darauf verpflichtete, sowohl die Quellen der Daten korrekt anzugeben als auch bei deren Vernetzung mit anderen Quellen auf eine Erhaltung der Urheberrechte zu achten. Gerade letzteres wurde aber offenbar als hinderlich für die Forschungspraxis empfunden, die unter enormen zeitlichem Druck stattfand. „The

dabate over GISAID and its governance exemplifies how efforts to abide by the principle of openness, particularly when openness is interpreted first and foremost as a form of sharing, can clash with responsible research measures geared towards protecting researchers whose work is unrecognized and/or discriminated against" (Leonelli 2023: 22).

Ein zweites Beispiel betrifft den Bereich *Open Infrastructure*. Konkret geht es um die Nutzung von Open-Source-Software durch Wissenschaftler:innen der Dritten Welt. Produkte dieser Art sind ursprünglich gedacht, um Personengruppen leistungsfähige IT-Arbeitswerkzeuge zur Verfügung zu stellen, die konkurrenzfähig mit hochpreisigen kommerziellen Lösungen sind, deren Nutzung für Mitarbeitende in finanzschwachen Einrichtungen meist nicht möglich ist. Entgegen dieser Zielsetzung zeigt nun eine Studie der *Global Young Academy*, auf welche sich Leonelli beruft, dass Forscher:innen in Ländern mit finanzschwächeren Institutionen gerade nicht zu Open-Source-Lösungen greifen, da sie durch deren Nutzung eine Stigmatisierung ihrer Arbeit befürchten. Leonelli erläutert diese Zwickmühle folgendermaßen: „One reason for this preference [die Nutzung kommerzieller Software-Produkte, NM] was the perceived stigma attached to using open software. Some participants in the study thought that editors and referees of international journals would interpret FOSS [free open-source software, NM] use as a mark of low-quality research, particularly when coming from research locations with little international reputations. Using well-recognized proprietary software such as *MathLab* and *Mathematica*, by contrast, was seen to align with global expectations around appropriate methodology, thus facilitating publications of research results by Anglo-American journals" (Leonelli 2023: 23). Im Raum steht die Frage danach, was als *gute* wissenschaftliche Methodik angesehen wird – und dazu zählen eben auch entsprechende Softwareprodukte. Offenbar besteht die Befürchtung, dass eine Abweichung von der Norm durch die Nutzung von OS-Produkten als Hinweis auf eine geringere epistemische Kompetenz der Forscher:innen gewertet werden würde.

Hier besteht tatsächlich eine Analogie hinsichtlich der Nutzung von OS-Produkten und -Methodiken in der Ersten Welt. Insbesondere der akademische Nachwuchs scheint einen schweren Stand zu haben, sich von etablierten Produkten – z.B. namenhaften Journals – in den Fächern bewusst zu distanzieren und auf OS-Alternativen zu setzen.[19] Auch in diesen Fällen besteht eine der Hauptsorgen darin, dass die eigene Forschung von Mitgliedern der eigenen Community als qualitativ minderwertig eingestuft werden könnte.

Aus diesen Überlegungen lässt sich folgern, dass die erfolgreiche Implementierung von OS-Vorhaben von einer Reihe von Faktoren abhängt und nicht allein durch das Teilen von Ressourcen realisiert werden kann. Diese Schlussfolgerung zieht auch Leonelli. Die unterschiedlichen *Quellen epistemischer Diversität* (vgl. Leonelli 2022) können begrifflicher, materieller, methodologischer, infrastruktureller, sozio-kultureller und auch institutioneller Art sein. Entscheidend sei, das eine Veränderung in einem dieser Bereiche Auswirkungen auf den Wissensgewinn und das wissenschaftliche Verstehen eines Phänomens haben könne.

Darüber hinaus betont die Autorin, dass diese Einflussfaktoren nicht anhand klassischer akademischer Disziplinengrenzen festgemacht werden können. Auch innerhalb eines einzelnen Faches – z.B. der Teilchenphysik – können unterschiedliche Standards im konkreten Arbeitsalltag der Wissenschaftler:innen zum Einsatz kommen. Wichtiger sei die Ebene von Arbeitsgruppen, innerhalb derer Festlegungen auf bestimmte Methodiken, Begriffe etc. erfolgten. Geht es um die Realisierung von OS-Initiativen müsse diesem Umstand Rechnung getragen werden, indem geplante Maßnahmen sich eher an solchen kleineren Einheiten orientierten, statt akademische Disziplinen als Ganzes in den Blick nehmen zu wollen.

Für Leonelli ergibt sich aus diesen Überlegungen die Forderung, das OS-Vorhaben als solches grundlegend konzeptionell neu zu fokussieren. „[…] I propose to invert this conceptualization of the direction of travel for OS implementation, which […] sees efforts to make research transparent as the starting point to

improve its quality and, eventually, its capacity to include. Instead, the implemantation of OS needs to start from consideration of what it may take to make research more inclusive, diverse and just […]" (Leonelli 2023: 42). Um diese Neuausrichtung zu erreichen, sei es notwendig, OS-Initiativen *prozess-* statt wie bisher primär *produktorientiert* aufzusetzen. Das bedeutet, dass Vertreter:innen von OS-Initiativen nicht mehr hauptsächlich vom Gedanken des Teilens von Ressourcen (Daten, IT-Produkte etc.) ausgehen sollten. Stattdessen sollten sie bestehende epistemische Praxen in ihren jeweiligen Dynamiken und Komplexitäten lokaler Arbeitsgruppen etc. in den Blick nehmen und mit deren Angehörigen in engen kommunikativen Austausch treten, um die integrative Zielsetzung von OS vorantreiben zu können. „[…] I propose a philosophy of openness predicated on a process-oriented view, whereby research is understood first and foremost as an effort to foster collective agency, grounded on intimate forma of relationality and trust, among widely diverse individuals and groups […]" (Leonelli 2023: 43).

Diese Forderung basiert auf der Einsicht, dass bestehende Best-Practice-Lösungen innerhalb erfolgreich funktionierender Arbeitsgruppen immer auch als Demarkationskriterien angesehen werden. Sprich, die Verwendung bestimmter Methoden, die Wahl einschlägiger Publikationsorgane, die Nutzung von Datenbanken spezifischer Anbieter etc. werden von den Wissenschaftler:innen verstanden als ein Kriterium für die Qualität der erzielten wissenschaftlichen Leistung (ggf. sogar für die Unterscheidung von wissenschaftlichen und pseudowissenschaftlichen Arbeitsweisen). OS-Initiativen müssen diesem Umstand Rechnung tragen, wenn sie mit ihren Zielsetzungen langfristig erfolgreich sein wollen, so die Diagnose von Leonelli.

Muss als Konsequenz also festgehalten werden, dass OS-Initiativen in der derzeitigen Form grundsätzlich als kontraproduktiv anzusehen sind? Dass sie nicht zur Lösung der Herausforderungen und Schwierigkeiten für die digitale Wissenschaftspraxis dienen können? Dass sie, im Gegenteil, eine ganze Reihe weiterer schwerwiegender Probleme für die Wissenschaft und damit auch für die Wissenschaft

als Lieferantin gesellschaftlich relevanter Informationen generiert? Nicht unbedingt.

3. Notwendigkeit der Neukonzeption von Open Science?

Zwei Punkte seien vorab klargestellt: (1) OS-Vorhaben sind in der Tat mit einer Vielzahl von aktuellen Herausforderungen konfrontiert – in Institutionen der Ersten Welt ebenso wie in der Dritten. Die vorgestellten empirischen Studien, auf die sich Leonellis kritische Analyse stützen, zeigen dies deutlich. (2) Es ist nicht sinnvoll, diese Herausforderungen einfach zu ignorieren, will man den gesteckten Zielsetzungen von OS gerecht werden. Hierzu zählen insbesondere die Hinweise auf die Relevanz epistemischer Diversität sowie auf die Betrachtung bestehender Arbeitspraxen als Evaluationskriterien für die angebliche Qualität wissenschaftlicher Leistungen.

Dennoch scheint es fraglich, ob die Konstatierung bestehender Schwierigkeiten zu einer so grundlegenden Neukonzeptionierung von OS-Vorhaben führen muss, wie Leonelli dies in ihren Texten einfordert. Im Folgenden soll auf der Basis derselben Problemstellungen eine alternative Lesart vorgestellt werden, welche die Fehlersuche aus den Zielsetzungen und Entwicklungsstrategien der OS-Initiative zurück in die traditionelle wissenschaftliche Gemeinschaft und deren Arbeitspraxis und -verständnis verlagert. Hieraus werden zwei sich ergänzende Verbesserungsvorschläge abgeleitet, die bei der Realisierung von OS-Vorhaben berücksichtigt werden sollten.

3.1 Eine Frage der Wissenschaftskultur

Leonelli konstatiert zu Recht, dass die Frage nach einer erfolgreichen Etablierung von OS immer auch verknüpft ist mit Überlegungen innerhalb der wissenschaftlichen Community nach Qualitäts- und damit nach Abgrenzungskriterien. „Systems of

research practice are systems of demarcation and exclusion. By setting criteria for what constitutes proper science and what does not, and which forms of expertise are deemed to be relevant, demarcation strategies provide the glue that brings and keeps epistemic activities together […]" (Leonelli 2023: 35). Wissenschaftler:innen definieren aus den ihnen bekannten Praxen, wie gute wissenschaftliche Arbeitsweisen aussehen sollten. Aus dem Ist-Zustand wird eine normative Forderung abgeleitet, ein klassischer Fehlschluss, wie man in der Philosophie feststellen würde. Gleichermaßen schwierig erscheinen die von Leonelli gezogenen Schlussfolgerungen aus dieser fehlerbehafteten Ausgangslage. Das Hauptproblem dabei liegt in der Angleichung an falsche Standards.

Letztlich zeigen die von ihr diskutierten Beispiele, dass der ausschlaggebende Faktor, an welchem die Realisierung von OS-Vorhaben scheitert, die Wissenschaftskultur in (außer-)universitären Forschungseinrichtungen der westlichen Welt darstellt. Hier scheint nach wie vor eine gelebte OS-Kultur zu fehlen. OS-Vorhaben im Bereich Open Infrastructure scheitern in der Dritten Welt daran, dass die oft berechtigte Befürchtung besteht, dass Wissenschaftler:innen in der Ersten Welt die Nutzung von OS-Werkzeugen mit minderwertiger Forschung assoziieren.

Dasselbe Problem wurde für den wissenschaftlichen Nachwuchs in der westlichen Welt angesprochen. Eine bewusste Entscheidung, alternative Technologien, Methoden oder Publikationsformen zu verwenden, scheitert an den eingefahrenen Arbeitspraxen etablierter Wissenschaftler:innen. Auch hier ist es v.a. das *Reputationsdenken*, das Veränderungen hin zu OS behindert.[20] Hierzu zählt einerseits die Annahme, dass ein namenhafter Publikationsort (d.h. eine entsprechende Fachzeitschrift im Falle von Artikeln oder ein Wissenschaftsverlag bei Monographien) ein aussagekräftiger Indikator für die vermeintlich hohe Qualität der dort veröffentlichten Ergebnisse sei. Andererseits werden Qualitätsannahmen daran geknüpft, ob eine Publikation in einer der großen Datenbanken (Scopus, Web of Science, Dimensions) indexiert wurde. Doch weder die erste noch die zweite Annahme bezüglich

vermeintlicher Qualitätsindikatoren ist haltbar. Schließlich werden Metriken zur Beurteilung vermeintlicher wissenschaftlicher Expertise herangezogen, die auf der Basis der Indexierungen in den genannten Datenbanken erhoben werden. Gemeint sind damit v.a. der sogenannte „Impact Factor" als auch der „h-Index". Beide Werte bauen auf der Annahme auf, dass die Anzahl von Zitaten der (indexierten) Artikel etwas über deren Relevanz für die wissenschaftliche Community und damit über die vermeintliche Expertise des Autors / der Autorin aussage (vgl. Deutsche Forschungsgemeinschaft, AG Publikationswesen 2022: 34ff.; Andersen 2019; Retzlaff 2022). Es kann jedoch leicht gezeigt werden, dass diese Metriken diesen Aussagewert nicht haben (vgl. z.B. Mößner 2022b).

Einerseits besteht die Möglichkeit, die vermeintlich objektiven Zahlenwerte zu manipulieren, z.B. durch die Bildung sogenannter Zitationskartelle, in welchen *Gefälligkeitszitate* vergeben werden. Andererseits kann die zu Grunde gelegte Datenbasis für die Erstellung der Metriken grundsätzlich hinterfragt werden. Hanne Andersen hält entsprechend fest: „Sometimes the number of citations of a publication is seen as a measure of the publication's quality. However, scientific results can be of a very high quality without necessarily gaining a high number of citations" (Andersen 2019: 151).

Quantitative Metriken und das Streben nach ihrer Optimierung wird von Expert:innen schon seit Längerem kritisch gesehen. So schreibt die DFG, dass es „nicht zielführend [sei], wenn Metriken über ihre Funktion als Orientierungshilfe hinaus bei der Einschätzung einer wissenschaftlichen Leistung dominieren oder sogar die alleinige Basis für eine solche Einschätzung darstellen" (Deutsche Forschungsgemeinschaft, AG Publikationswesen 2022: 35). Trotz aller berechtigter Kritik hat sich aber die bibliometrische Bewertungspraxis tief in die Wissenschaftskultur eingegraben.

Um hier Abhilfe zu schaffen, wurde die europäischen Initiative „Coalition for Advancing Research Assessment (CoARA)" ins Leben gerufen. Forschungseinrichtungen sind aufgerufen, sich durch eine Unterzeichnung der Erklärung auf die formulierten

Zielsetzungen für eine Reform der Bewertungspraxis wissenschaftlicher Leistung zu verpflichten. In den Leitlinien heißt es u.a.: „Signatories commit to a common vision, which is that the assessment of research, researchers and research organisations recognises the diverse outputs, practices and activities that maximise the quality and impact of research. This requires basing assessment primarily on qualitative judgement, for which peer-review is central, supported by responsible use of quantitative indicators."[21] Zu den Unterzeichnenden zählt beispielsweise die DFG, die sich damit um eine Abkehr von quantitativen Metriken und um eine verstärkte Hinwendung zu einer qualitativen Beurteilung wissenschaftlicher Leistungen bemüht (Deutsche Forschungsgemeinschaft 2022: 11).[22]

Dies ist nur ein Beispiel dafür, was sich innerhalb der bestehenden Wissenschaftskultur der westlichen Länder ändern müsste, um OS-Vorhaben künftig erfolgreicher voranbringen zu können.[23] Leonellis Vorschlag zöge genau in die entgegengesetzte Richtung, sprich, zu einer Angleichung der Grundkonzeption von OS-Initiativen an bestehende fehlerbehaftete Wissenschaftspraxen der westlichen Welt. Um diese Fehlentwicklung zu vermeiden, besteht ein erster wichtiger Schritt in einer kritischen Reflexion der eigenen Praxis – die, wie Leonelli richtig bemerkt, von den Beteiligten oft als Gold-Standard angesehen wird, wenn es darum geht zu definieren, was eine gute wissenschaftliche Arbeitsweise ausmache.

Zu dieser Reflexion gehören Fragen danach, (1) wie wissenschaftliche Leistung tatsächlich evaluiert werden sollte – Initiativen wie CoARA bieten Ansatzpunkte für Verbesserungen; (2) was unter ‚wissenschaftlicher Expertise' verstanden wird – denn nur so lassen sich überhaupt sinnvolle Kriterien für deren Bewertung aufstellen; (3) was als Zielsetzung wissenschaftlicher Arbeit angesehen wird – insbesondere im Sinne einer Abwägung, ob die Suche nach Wahrheit im Vordergrund stehen soll oder die Entwicklung marktfähiger Produkte.

Beachtet werden muss ferner, dass das Wissenschaftssystem ein stark hierarchisches Gefüge ist. Hier Änderungen durchzusetzen bedeutet, dass man bestehende Strukturen berücksichtigen muss.

Um OS-Vorhaben voranzutreiben, braucht es zum einen *Zugpferde*, d.h. etablierte Wissenschaftler:innen oder Institutionen, welche die Veränderungen unterstützen und die neue Kultur vorleben. Eine Möglichkeit in diesem Kontext bestünde z.B. darin, dass Fachgesellschaften bei ihren Publikationen auf Open-Access-Strategien nach dem Diamond-Modell setzen.[24] Zum anderen muss der wissenschaftliche Nachwuchs rechtzeitig für Probleme im Zusammenhang mit der Praxis des elektronischen Publizierens und der Verwendung von Metriken zur Wissenschaftsbeurteilung sensibilisiert werden. Ziel muss es sein, künftige Forscher:innen über bestehende Schwierigkeiten aufzuklären und ihnen die Möglichkeit an die Hand zu geben, sich für Alternativen zu entscheiden bzw. diese selbst aktiv mitzugestalten.

3.2 Re-Lokalisierung wissenschaftlicher Infrastruktur

Damit zusammenhängend wird deutlich, dass die Forderung naheliegend ist, die relevante (IT-)Infrastruktur zurück in die Hände der wissenschaftlichen Community zu legen. Diesen Ansatz hat u.a. Björn Brembs, Mitinitiator der Initiative „Stop Tracking Science", prominent hervorgehoben. Im Hinblick auf das Problem des Datentrackings durch die großen Wissenschaftsverlage plädieren Brembs et al. für ein Ersetzen der „traditional and outdated academic journals with an open, interoperable, and community-governed modern scholarly information infrastructure" (Brembs u.a. 2023: 12). Nur auf diese Weise könne die Monopolstellung der Informationsanalysten gebrochen werden, die zu einer Reihe von Fehlentwicklungen im Bereich des elektronischen Publizierens und der Wissenschaftsevaluation geführt haben.

Ein solches communitybasiertes Vorhaben wird in Diamond-OA-Modellen angestrebt.[25] Für den Erfolg dieser alternativen Publikationsmodelle gilt, dass sie nicht zu einer Mehrbelastung der Wissenschaftler:innen im Alltagsgeschäft führen dürfen. Sprich, auch wenn es heißt, dass die Infrastruktur *zurück in die Hände der Community* gelegt werden solle, bedeutet das nicht,

dass Forscher:innen nun selbst alle notwendigen Arbeitsschritte für die Publikation, Archivierung und Auffindbarmachung ihrer Beiträge übernehmen sollen. An dieser Stelle muss man sich vergegenwärtigen, dass Forschungseinrichtungen per se über eine Reihe von Infrastruktureinrichtungen verfügen, mit denen eine engere Zusammenarbeit in OS-Vorhaben realisiert werden könnte. Dazu zählen in erster Linie Bibliotheken[26] als auch Rechenzentren.

Ein Beispiel sei zu diesem Komplex näher ausgeführt: Leonellis Analyse (2023) hatte aufgezeigt, dass eine Hürde für die Realisierung von OS-Vorhaben für marginalisierte Gruppen oft im Zusammenhang mit der berechtigten Sorge vor Benachteiligungen bestehe, z.B. der Ausnutzung von Daten, die durch Mitglieder dieser Gruppen erhoben und bereitgestellt wurden, ohne dies nachfolgend adäquat wertschätzend zu attribuieren. Kurz, es besteht die Angst vor *Datenklau* durch Forschende der Ersten Welt.

Um diesem Problem vorzubeugen, wurde innerhalb einer solchen Gruppe nun selbst die Initiative ergriffen. Konkret geht es um das Bibliothekssystem *Koha*,[27] eine Open-Source-Lösung, wie sie von Leonelli in ihrem Beitrag kritisiert wurde. Dieses System zeichnet sich dadurch aus, dass es von Mitgliedern einer marginalisierten Gruppe – nämlich den Māori in Neuseeland – für die Nutzung durch andere Gruppen, u.a. der Aborigines in Australien, entwickelt wurde.[28] Bei dieser Entwicklung stand gerade die oben ausgedrückte Sorge im Vordergrund. Folglich bietet dieses System die Möglichkeit, die Oberhoheit über die eigenen Daten zu behalten, die in das System eingespeist werden. Nutzer:innen der Anwendung können selbst entscheiden, ob und, wenn ja, wie ihre Daten weiter verwendet werden dürfen. Die Open-Source-Lösung grenzt sich damit klar von klassischen kommerziellen Produkten ab, die eine solche Entscheidungsmacht der Nutzer:innen typischerweise unterbinden. *Koha* stellt damit eine OS-Lösung vor, die den ursprünglichen Zielsetzungen von OS-Initiativen entspricht, ohne von den negativen Effekten der bestehenden westlichen Wissenschaftskultur verfälscht worden zu sein. Es zeigt sich also, dass

die Zielsetzungen der Erhöhung von Diversität und Inklusion durchaus nicht dazu führen müssen, OS-Initiativen grundlegend neu zu konzeptionieren, wie Leonelli es anregt. Auch bei *Koha* geht es letztlich um das von ihr kritisierte *Teilen* von Daten.

4. Resümee

Zusammenfassend lässt sich feststellen, dass OS-Initiativen mehrdimensional auf unterschiedliche Ziele ausgerichtet sind. Nicht alle dieser Ziele lassen sich gleichermaßen schnell und einfach realisieren. Einige – insbesondere solche, die auf eine Erhöhung von epistemischer Diversität und eine breitere Inklusion abzielen – bedürfen Änderungen, die sicherlich nicht allein mit der Bereitstellung technologischer Lösungen erreicht werden können.

Deutlich wurde, dass ein entscheidender Faktor in der Entwicklung von OS-Vorhaben in der *Wissenschaftskultur* besteht. OS hat viel mit dem Gedanken der *Autonomie der Wissenschaft* zu tun. Gegen diese steht derzeit vor allem die stetig wachsende Kommerzialisierung dieses Sektors mit der immer stärker werdenden Abhängigkeit wissenschaftlicher Arbeitsprozesse von Unternehmensprodukten. Neoliberale Ansätze führen oft dazu, dass deren Anhänger in allem einen Markt sehen wollen, wie Wendy Brown (2015) kritisch anmerkt, aber nicht überall ist der Marktgedanke sinnvoll – v.a. dann nicht, wenn Kosten sozialisiert und Profite privatisiert werden, wie es häufig im Wissenschaftssektor zu beobachten ist. Dies ist eine wichtige Einsicht, welche im Hinblick auf die Forderung nach einer Änderung der westlichen Wissenschaftskultur berücksichtigt werden sollte.

Eine solche kulturelle Neustrukturierung ist ein längerfristiges Unterfangen, das Vorbilder notwendig macht. Etablierte Wissenschaftler:innen und Fachgesellschaften stellen Akteure dar, die eine entsprechende Vorreiterfunktion übernehmen können.

Zu den Aspekten, die im Hinblick auf einen kulturellen Wandel in den Wissenschaften zu beachten sind, zählen neben der kritischen Hinterfragung kommerzieller Praxen auch Überlegungen

dazu, was innerhalb der Community als das Ziel wissenschaftlicher Aktivität betrachtet wird und, damit zusammenhängend, was eine entsprechende wissenschaftliche Expertise auszeichnet.

Darüber hinaus ist es korrekt, dass zum Erfolg von OS-Vorhaben auch eine funktionierende (technologische) Infrastruktur an den verschiedenen Einrichtungen sowie übergreifend zählt. Diese Infrastruktur einzurichten und langfristig vorzuhalten darf dabei allerdings nicht zu einer Zusatzbelastung für die Wissenschaftler:innen werden. Hier ist eine kompetente Unterstützung durch entsprechende Einrichtungen wie Bibliotheken etc. erforderlich. Durch diese Neupositionierung können sie im Gefüge der Forschungslandschaft auch eine neue Stärkung erfahren. Um sie in diesem Zusammenhang sinnvoll zu unterstützen, muss auch eine Erörterung der Frage nach anfallenden Kosten, z.B. wenn es um die Langzeitarchivierung von Forschungsdaten geht, zwischen den beteiligten Akteur:innen jenseits von kurzfristigen Projektförderungen erfolgen.

Anmerkungen

[1] Zu diesem Punkt vgl. auch den Beitrag von Bernhard Mittermaier in diesem Band.

[2] Vgl. die Selbstdarstellungen unter https://www.elsevier.com/about, https://www.wiley-vch.de/de/ueber-wiley/wiley-weltweit, https://group.springernature.com/de/group/about-us, eingesehen am 22. Juli 2024. Vgl. auch die Analyse von Ulrich Herb (2018), der festhält, dass Elsevier spätestens seit 2004 „nicht mehr nur ein Verlag" sei, da für das Unternehmen mit der Marktetablierung der firmeneigenen Datenbank *Scopus* Geschäftsmodelle im Bereich der Bibliometrie eröffnet wurden. Herb dokumentiert in seinem Beitrag ausführlich, in welchen Bereichen des wissenschaftlichen Arbeitsprozesses die Produkte von Elsevier zwischenzeitlich Verwendung finden und wie dadurch ein umfassendes Profiling eines jeden Einzelindividuums für das Unternehmen möglich wird.

[3] Vgl. dazu die Übersichtsgrafik der Plattformbetreiber in den unterschiedlichen Arbeitsbereichen der Wissenschaftler:innen unter https://stoptrackingscience.eu/background-information/, eingesehen am 22. Juli 2024.

[4] RELX steht für Reed, Elsevier und LexisNexis.
[5] Vgl. dazu auch die Beiträge in (Mößner und Erlach 2022).
[6] Eine Übersicht zu laufenden Projekten dieser Art in Deutschland findet sich unter z.B. unter https://www.mitforschen.org/, eingesehen am 23. Juli 2024 sowie international unter https://www.zooniverse.org/, eingesehen am 23. Juli 2024. Bekannt sind auch die vielfältigen Projekte der NASA im Bereich der citizen science, vgl. https://science.nasa.gov/citizen-science/, eingesehen am 23. Juli 2024.
[7] Vgl. https://zoo4.galaxyzoo.org/, eingesehen am 23. Juli 2024.
[8] Gemeint sind hiermit z.B. Daten, Modelle, biologische Proben und anderes mehr je nachdem, was in der betreffenden akademischen Disziplin als Grundlage für Forschungsprozesse benötigt wird.
[9] Vgl. Gehring 2024.
[10] „Re-Kalibrierung – Wissenschaftskommunikation im digitalen Wandel", Württembergische Landesbibliothek Stuttgart 2022.
[11] Auf diesen Punkt weist Evgeny Morozov (2015) hin.
[12] Ein ähnliches Argument diskutiert im Übrigen Edward Snowden (2019) im Kontext der Überwachungsaktivitäten durch die amerikanischen Geheimdienste.
[13] Verstanden werden hierunter technologische Systeme, die intellektuellen Leistungen des Menschen gleichkommen.
[14] An dieser Stelle muss festgehalten werden, dass die nachfolgend beschriebenen Prozesse insofern nicht dem von Floridi erläuterten gängigen Mechanismus des *Envelopings* folgen, als letzteren eine bewusste Entscheidung zu einer Beförderung der Effizienzsteigerung schwacher KI-Systeme vorausgeht. Dies ist jedoch nicht der Fall, wenn hier vom Training generativer KI-Systeme anhand von freiverfügbar gemachten Daten und Texten gesprochen wird.
[15] Es ließe sich fragen, inwiefern eine solche Nutzung von OA-Texten durch die für sie vergebenen Lizenzen unterbunden werden können.
[16] Auf ein Folgeproblem sei hingewiesen: Wenn Schüler:innen und Studierende von Anfang an mit diesen Technologien arbeiten, erwerben sie nicht mehr die Fähigkeiten, die ihnen eine kritische Überprüfung der von der KI generierten Ergebnisse erlauben würde. Wenn diese Fähigkeiten aber nicht vorhanden oder in entscheidendem Maße geschwächt sind, kann davon ausgegangen werden, dass die Menge an falschen Informationen, die durch KI-generierte Texte in Umlauf gebracht werden, exponentiell steigen wird. Es fehlt ja eine entsprechende Korrekturinstanz. Man kann diesen Gedanken noch weiterspinnen, indem man sich überlegt, dass die Datenbasis, aus der die

KI lernt, damit ebenfalls mehr und mehr Fehler enthält, wie auch im nachfolgenden Punkt oben angesprochen.

[17] Zusammengefasst in (Leonelli 2023).

[18] Gemeint sind hiermit nicht allein Mitglieder der schon genannten Gruppen, sondern auch Bürger:innen der westlichen Welt. Konzepte der sogenannten *citizen science* tragen diesem Anspruch Rechnung, wie oben ausgeführt wurde.

[19] Auf diesen Punkt weist eine Reihe von Expert:innen der Informationswissenschaften und auch der Forschungsförderungsinstitutionen wie der DFG hin (vgl. z.B. die entsprechenden Beiträge in Mößner und Erlach 2022).

[20] Frank Miedema (2022) schildert in seinem Praxisbericht über die UMC Utrecht, inwiefern dieses Reputationsdenken sich nachteilig auf Entwicklungen im Bereich der angewandten Wissenschaften auswirkt und inwieweit sich daraus eine Initiative für konkrete OS-Vorhaben an seiner Hochschule ergeben hat: „In fact, we also realized that our research evaluation system did not acknowledge top professionals and clinicians engaged in more practical patient-driven research where journal impact factors are lower and no prestigious personal career grants are to be won" (ebd.: 164f.).

[21] https://coara.eu/agreement/the-agreement-full-text/, eingesehen am 02. September 2024.

[22] Zur CoARA-Initiative vgl. https://coara.eu/, eingesehen am 28. August 2024.

[23] Inwiefern das an quantitative Metriken gebundene Reputationsdenken in der wissenschaftlichen Gemeinschaft als nachhaltiger Störfaktor auftritt und dabei die Marktmacht der Informationsanalysten weiter stützt, wird im Positionspapier „Elektronisches Publizieren und Bewertung wissenschaftlicher Expertise" ausführlich erläutert (vgl. Mößner und Erlach 2024 sowie im Anhang dieses Bandes).

[24] Vgl. dazu auch den Beitrag „Diamond Open Access?" von Stefan Drößler in diesem Band.

[25] Es sei angemerkt, dass auch diese OS-Vorhaben nicht frei von Problemen sind. Insbesondere stellt sich häufig die Frage einer nachhaltigen Finanzierung der oftmals auf Projektebene begonnenen Initiativen.

[26] Vgl. dazu wiederum das Projekt *Diamond_Thinking* der Universitätsbibliothek Stuttgart, vorgestellt durch Stefan Drößler in diesem Band.

[27] Vgl. Koha Library Software https://koha-community.org/, eingesehen am 04. September 2024.

[28] Vgl. dazu Chris Cormack (2023) Beitrag „Libraries and Indigenious data sovereignity" https://perlkohacon.fi/Schedule.html#?talk_id=32, eingesehen am 04. September 2024.

Literatur

Andersen, Hanne (2019) „Can Scientific Knowledge Be Measured by Numbers?" In *What Is Scientific Knowledge?: An Introduction to Contemporary Epistemology of Science*, hrsg. von Kevin McCain und Kostas Kampourakis. Routledge, S. 144–159. https://doi.org/10.4324/9780203703809-10.

Brembs, Björn, Philippe Huneman, Felix Schönbrodt, Gustav Nilsonne, Toma Susi, Renke Siems, Pandelis Perakakis, Varvara Trachana, Lai Ma, und Sara Rodriguez-Cuadrado (2023) „Replacing Academic Journals". *Zenodo*. https://doi.org/10.5281/ZENODO.7974116.

Brown, Wendy (2015) *Undoing the Demos. Neoliberalism's Stealth Revolution*. Zone Books.

Deutsche Forschungsgemeinschaft (2022) „Open Science als Teil der Wissenschaftskultur. Positionierung der Deutschen Forschungsgemeinschaft". *Zenodo*. https://doi.org/10.5281/ZENODO.7193838.

Deutsche Forschungsgemeinschaft, AG Publikationswesen (2022) „Wissenschaftliches Publizieren als Grundlage und Gestaltungsfeld der Wissenschaftsbewertung". *Zenodo*. https://doi.org/10.5281/ZENODO.6538163.

DFG-Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme (2021) „Datentracking in der Wissenschaft: Aggregation und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage. Ein Informationspapier des Ausschusses für Wissenschaftliche Bibliotheken und Informationssysteme der Deutschen Forschungsgemeinschaft". *Zenodo*. https://doi.org/10.5281/ZENODO.5900759.

Elliott, Kevin C. (2022) „A Taxonomy of Transparency in Science". *Canadian Journal of Philosophy* 52 (3): 342–355. https://doi.org/10.1017/can.2020.21.

Floridi, Luciano (2014) *The 4th Revolution. How the Infosphere Is Reshaping Human Reality.* Oxford University Press.

Gehring, Petra (2023) „Transparenz, Erklärbarkeit, Interpretierbarkeit, Vertrauen: digitalethische Doppelgänger des Verantwortungsbegriffs". *Deutsche Zeitschrift Für Philosophie* 71 (5): 629–645. https://doi.org/10.1515/dzph-2023-0051.

Gehring, Petra (2024) „As open as possible? Datenzugänge für die Wissenschaft im Spannungsfeld zwischen Open Science, Datenteilen und Datenkauf". In *Datenzugangsregeln – zwischen Freigabe und Kontrolle,* hrsg. von Steffen Augsberg, Marcus Düwell und Benjamin Müller, Campus, S. 77–100.

Herb, Ulrich (2018) „Zwangsehen und Bastarde. Wohin steuert Big Data die Wissenschaft?" *Information – Wissenschaft & Praxis* 69 (2–3): 81–88. https://doi.org/10.1515/iwp-2018-0021.

Lamdan, Sarah (2023) *Data cartels: the companies that control and monopolize our information*. Stanford University Press.

Leonelli, Sabina (2022) „Open Science and Epistemic Diversity: Friends or Foes?" *Philosophy of Science* 89 (5): 991–1001. https://doi.org/10.1017/psa.2022.45.

Leonelli, Sabina (2023) *Philosophy of Open Science*. Cambridge University Press. https://doi.org/10.1017/9781009416368.

Leonelli, Sabina, Rebecca Lovell, Benedict Wheeler, Lora Fleming und Hywel Williams (2021) „From FAIR Data to Fair Data Use: Methodological Data Fairness in Health-Related Social Media Research". *Big Data & Society* 8 (1): 205395172110103. https://doi.org/10.1177/20539517211010310.

Levin, Nadine, Sabina Leonelli, Dagmara Weckowska, David Castle und John Dupré (2016) „How Do Scientists Define Openness? Exploring the Relationship Between Open Science Policies and Research Practice". *Bulletin of Science, Technology & Society* 36 (2): 128–141. https://doi.org/10.1177/0270467616668760.

Longino, Helen E. (2013) „Werte, Heuristiken und die Politik des Wissens". In *Werte in den Wissenschaften. Neue Ansätze zum Werturteilsstreit*, hrsg. von Gerhard Schurz und Martin Carrier. Suhrkamp, S. 209–232.

Miedema, Frank (2022) *Open Science: The Very Idea*. Dordrecht: Springer. https://doi.org/10.1007/978-94-024-2115-6.

Morozov, Evgeny (2014) *To Save Everything, Click Here. Technology, Solutionism and the Urge to Fix Problems That don't Exist*. Penguin Books.

Morozov, Evgeny (2015) „Ich habe doch nichts zu verbergen". *APuZ Aus Politik und Zeitgeschichte* 65 (11–12): 3–7.

Mößner, Nicola (2022a) „›Keine Spyware in Bibliotheken und Institutionen‹ – wirklich?: Die Gefahren des Wissenschaftstrackings". *Mitteilungen der DGPhil* 55 (Frühjahr): 11–13.

Mößner, Nicola (2022b) „Vermessene Expertise". *OPUS*. https://doi.org/10.18419/OPUS-12157.

Mößner, Nicola (2023) „Noch nicht ist schon zu viel". *Jahrbuch Technikphilosophie Faktor Mensch*, Vol. 9, 253–253. https://doi.org/10.5771/9783748941767-253.

Mößner, Nicola und Klaus Erlach (Hrsg.) (2022) *Kalibrierung der Wissenschaft: Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis*. Transcript Verlag. https://doi.org/10.14361/9783839462102.

Mößner, Nicola und Klaus Erlach (2024) „Elektronisches Publizieren und Bewertung wissenschaftlicher Expertise". *Zenodo.* https://doi.org/10.5281/ZENODO.10948471. (vgl. in diesem Band, Kap. 12).

Mößner, Nicola und Philip Kitcher (2017) „Knowledge, Democracy, and the Internet". *Minerva* 55 (1): 1–24. https://doi.org/10.1007/s11024-016-9310-0.

O'Neil, Cathy (2016) *Weapons of Math Destruction. how big data increases inequality and threatens democracy*. Penguin Books.

Retzlaff, Eric (2022) „Wer bewertet mit welchen Interessen wissenschaftliche Publikationen?: Eine Skizzierung des Einflusses kommerzieller Interessen auf die Forschungsoutput-Bewertung". In *Kalibrierung der Wissenschaft: Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis,* hrsg. von Nicola Mößner & Klaus Erlach. Transcript Verlag, S. 139–162. https://doi.org/10.14361/9783839462102-006.

Schlude, Antonia, Ulrike Mendel, Roland A. Stürz und Micha Fischer (2024) „Verbreitung und Akzeptanz generativer KI an Schulen und Hochschulen". *Bidt DE*. https://www.bidt.digital/publikation/verbreitung-und-akzeptanz-generativer-ki-an-schulen-und-hochschulen/.

Siems, Renke (2022) „Das Lesen der Anderen. Die Auswirkungen von User Tracking auf Bibliotheken". *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB*, 9(1): 1–25. https://doi.org/10.5282/O-BIB/5797.

Siems, Renke (2023) „‚Überwachen und Strafen' – Tracking und Kontrolle des Forschungszyklus". *ABI Technik* 43 (2): 86–95. https://doi.org/10.1515/abitech-2023-0016.

Snowden, Edward J. (2019) *Permanent Record*. Macmillan.

Staunton, Ciara, Carlos Andrés Barragán, Stefano Canali, Calvin Ho, Sabina Leonelli, Matthew Mayernik, Barbara Prainsack und

Ambroise Wonkham (2021) „Open Science, Data Sharing and Solidarity: Who Benefits?“ *History and Philosophy of the Life Sciences* 43 (4): 115. https://doi.org/10.1007/s40656-021-00468-6.
UNESCO (2021) „UNESCO Recommendation on Open Science“. *UNESCO*. https://doi.org/10.54677/MNMH8546.
UNESCO und Canadian National Commission for UNESCO (2022) „An introduction to the UNESCO Recommendation on Open Science“. *UNESCO*. https://doi.org/10.54677/XOIR1696.
Vallbracht, Christine (2024) „Der menschliche Anteil geht zusehends zurück“. *Forschung & Lehre*. https://www.forschung-und-lehre.de/forschung/der-menschliche-anteil-geht-zusehends-zurueck-6320.
Zuboff, Shoshana (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. First trade paperback edition. PublicAffairs.

Internetquellen

CoARA. https://coara.eu/agreement/the-agreement-full-text/, eingesehen am 02. September 2024.
Cormack, Chris (2023) Beitrag „Libraries and Indigenious data sovereignity.“ https://perlkohacon.fi/Schedule.html#?talk_id=32, eingesehen am 04. September 2024.
Elsevier. https://www.elsevier.com/about, eingesehen am 22. Juli 2024.
Galaxy Zoo. https://zoo4.galaxyzoo.org/, eingesehen am 23. Juli 2024.
Koha Library Software. https://koha-community.org/, eingesehen am 04. September 2024.
Mitforschen. https://www.mitforschen.org/, eingesehen am 23. Juli 2024.
NASA. https://science.nasa.gov/citizen-science/, eingesehen am 23. Juli 2024.
Springer Nature. https://group.springernature.com/de/group/about-us, eingesehen am 22. Juli 2024.
„Stop tracking science.“ https://stoptrackingscience.eu/background-information/, eingesehen am 22. Juli 2024.
Wiley. https://www.wiley-vch.de/de/ueber-wiley/wiley-weltweit, eingesehen am 22. Juli 2024.
Zooniverse. https://www.zooniverse.org/, eingesehen am 23. Juli 2024.

KAPITEL 12

Positionspapier: „Elektronisches Publizieren und Bewertung philosophischer Expertise“[1]

Nicola Mößner und Klaus Erlach

1. Ausgangslage: Vom elektronischen Publizieren zur Entmündigung der Wissenschaft

Informations- und Kommunikationstechnologien (IuK-Technologien) sind fester Bestandteil der wissenschaftlichen Praxis geworden. Während jedoch beispielsweise Praktiken des Einsatzes von Big-Data-Anwendungen in Forschungsprozessen, von sozialen Medien im Bereich der Wissenschaftskommunikation und von e-Learning- und Lernmanagement-Systemen im Bereich der Lehre eine kritische Reflexion erfahren, bleibt eine Reihe weiterer IuK-Technologien, die unmittelbar von WissenschaftlerInnen in den täglichen Arbeitsprozessen genutzt werden, bisher mehr oder weniger unbeachtet. Im Folgenden soll

Zitationsvorschlag für dieses Kapitel:
Mößner, N. und Erlach, K. 2026. Positionspapier: „Elektronisches Publizieren und Bewertung philosophischer Expertise“. In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* S. 353–366. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.o. License: CC BY 4.0

der Fokus auf Technologien im Umfeld der Abläufe beim elektronischen Publizieren, bei der Informationsrecherche und bei der Evaluation von Forschungsergebnissen gelegt werden.

Der Einfluss der IuK-Technologien auf Erkenntnisgewinnung und -verbreitung ist dabei weitreichend. Deren feste Einbindung in die täglichen Arbeitsprozesse hat die AnwenderInnen in eine oftmals unbemerkte Abhängigkeit von diesen technischen Produkten gebracht, aus der sie sich selbst ohne umfangreiche Umgestaltungen ihres (sozialen) Arbeitsumfelds nicht lösen können. Diesen Umstand machen sich wiederum kommerzielle IuK-Anbieter zunutze, sodass in wachsendem Maße die Autonomie der Wissenschaft z.B. im Hinblick auf die Setzung von Forschungszielen und Themenschwerpunkten in Gefahr gerät, wie auch die Deutsche Forschungsgemeinschaft (DFG) kürzlich in einem Positionspapier festgestellt hat (DFG 2022: 38ff.).

1.1 Elektronisches Publizieren und der ‚Markt' der Informationen

Eine zentrale Rolle im Zusammenhang mit den skizzierten Entwicklungen spielen die großen Wissenschaftsverlage, die sich – wie z.B. Elsevier, Springer Nature und Wiley – in ihrem Produktportfolio zu umfassenden Informationsdienstleistern weiterentwickelt haben. Der Markt des elektronischen Publizierens (inkl. der Bereitstellung, Zugänglichmachung, Archivierung und Bewertung von Informationen) wächst in allen Wissenschaftszweigen beständig an. Da die zahlenden Kunden dieser Informationsdienstleister in der Regel von der öffentlichen Hand getragene Einrichtungen sind, die ihre Inhalte kostenfrei zur Verfügung stellen, führt das Geschäftsmodell der kommerziellen Akteure mit Renditen von um die dreißig Prozent dazu, Profite zu privatisieren und Kosten zu sozialisieren.

Die Grafik veranschaulicht diesen Kritikpunkt, der von ExpertInnen der Informationswissenschaft in seinen einzelnen Bestandteilen bereits vorgebracht worden ist: Während einerseits immer

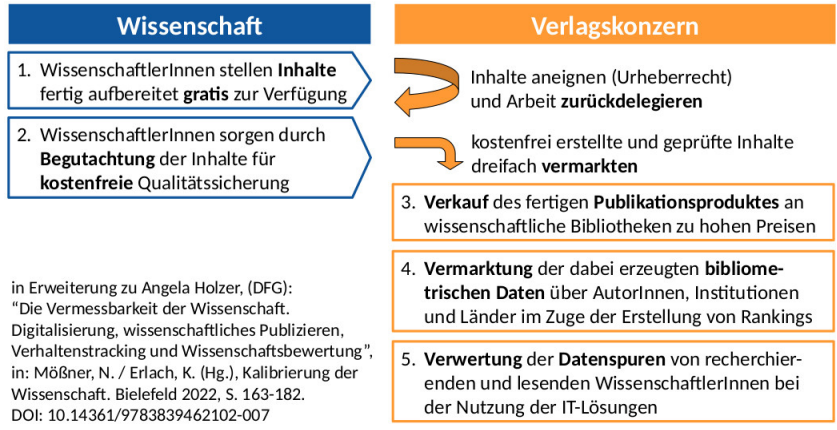


Abbildung 12.1: Arbeits- und Verwertungsschritte im Prozess des elektronischen Publizierens.

mehr originäre Verlagsaufgaben der Informationsaufbereitung (z.B. Erstellung eines Layouts, Übernahme eines Lektorats) und -bereitstellung (z.B. Begutachtungen) als kostenlose Dienstleistungen von den WissenschaftlerInnen selbst erbracht werden müssen, vermarkten andererseits die Informationsdienstleister auf einer wachsenden Anzahl von Aggregationsebenen innovative (Daten-)Produkte, die sie hochpreisig an diejenigen wissenschaftlichen Einrichtungen zurückverkaufen, die die Inhalte ursprünglich kostenfrei zur Verfügung gestellt haben.

1.2 Informationsdienstleister und die Metrisierung der Wissenschaft

Verlagsoligopole, wie im ersten Abschnitt beschrieben, erweisen sich derzeit als besonders erfolgreich im Bereich der Kundenbindung. Dieser Umstand lässt sich u.a. darauf zurückführen, dass die Bündelung der durch die WissenschaftlerInnen bereitgestellten Informationen in Datenbanken (wie Scopus, Web of Science, Dimensions) nicht allein ein effizientes Werkzeug der globalen Informationsrecherche darstellt, sondern darüber hinaus die

Basis der metrischen Beurteilung wissenschaftlicher Leistungen (ausgedrückt als h-Index, Impact Faktor usw.) bildet.

Die bekannten bibliometrischen Verfahren greifen durchgängig auf Zitationszahlen zurück, deren Erhebung nur effizient möglich ist, wenn ein entsprechendes digitales Data-Mining in wissenschaftlichen Publikationen durchgeführt werden kann. Datenbanken sind also unumgänglich, um den Bestand wissenschaftlicher Publikationen umfassend erschließen und nach den genannten Gesichtspunkten auswerten zu können. Es liegt im Eigeninteresse der Informationsdienstleister, diese Produkte in der wissenschaftlichen Praxis zu verankern, indem sie suggerieren, dass allein durch die Nutzung derartiger Datenbanken qualitativ hochwertige Informationen für die Arbeit der WissenschaftlerInnen verfügbar würden. Letztere sind wiederum nicht in der Lage, sich diesem System der Metrisierung zu entziehen, da es von hoher Relevanz für die Etablierung im Beruf sowie die wissenschaftliche Karriere ist.

1.3 Stärkung der Autonomie der Wissenschaften

In dieser Problem-Skizze wird deutlich, dass ein Ausstieg aus diesem technologie-induzierten Abhängigkeitsverhältnis der wissenschaftlichen Praxis von global agierenden Informationsdienstleistern erfordern würde, die Marktmacht der Oligopolisten zu brechen oder zu unterlaufen. Björn Brembs (vgl. Brembs et al. 2023) hat diesem Punkt mit der Forderung Ausdruck verliehen, dass entweder (a) ein echter Markt für entsprechende IT-Dienstleistungen geschaffen oder (b) die relevante (IT)-Infrastruktur zurück in die Hände der Bildungseinrichtungen geben werden müsste. Beides setzt allerdings ein zielgerichtetes, gestaltendes Eingreifen mehrerer starker Akteure der öffentlich-rechtlichen Ebene voraus.

Im Rahmen der DGPhil-Forumstagung „verTRACKTe Infrastruktur?! Von Konstrukteuren und Kontrolleuren wissenschaftlicher Expertise“, unterstützt vom IZKT der Universität Stuttgart, wurde dieser Vorschlag zum Anlass genommen, nach

Lösungsalternativen zu suchen, die unmittelbar aus der wissenschaftlichen Community heraus eigenständig umgesetzt werden können. Im interdisziplinär besetzten Workshop wurden dabei eine Reihe von Vorschlägen erarbeitet, die sich insbesondere an die akademische Philosophie als Fachdisziplin wenden und im Folgenden den VertreterInnen der Fachgesellschaften zur weiteren Diskussion und Umsetzung vorgestellt werden.

2. Empfehlungen zur Stärkung der Autonomie der Wissenschaften

Die nachfolgenden Empfehlungen betreffen Änderungen der Wissenschaftskultur innerhalb der akademischen Philosophie in Deutschland. Sie sind überwiegend lokal umsetzbar. Allerdings ist die erfolgreiche Realisierung einiger dieser Maßnahmen langfristig und nachhaltig nur dadurch sicherzustellen, dass auch Veränderungen im internationalen Kontext des Wissenschaftsgefüges angestrebt werden. Die vorliegenden Empfehlungen sind insofern als ein erster Schritt einer stufenweisen Veränderung zu denken.

Die Empfehlungen zielen auf eine Stärkung der Autonomie des Wissenschaftsbetriebs. Die einzelnen Maßnahmen sollen dazu beitragen, den Bereich des elektronischen Publizierens stärker im Sinne der wissenschaftlichen Community zu gestalten. Da, wie oben gezeigt, von der informationstechnologischen Grundlage ebenfalls die Frage nach dem Umgang mit der zunehmenden Metrisierung der Wissenschaftsbewertung abhängt, umfassen die folgenden Punkte auch Vorschläge zu Änderungen in diesem Bereich.

2.1 Elektronisches Publizieren: Infrastruktur in den Händen der wissenschaftlichen Community

Ziel des wissenschaftlichen Publizierens ist die zeitnahe Bereitstellung von Forschungsergebnissen und Fachwissen für die jeweilige

Fachcommunity sowie teilweise auch für andere Fachwissenschaften als Grundlagenwissenschaft. Publikationen für wissenschaftsexterne Zielgruppen (interessierte Laien, SchülerInnen etc.), wie sie in der Philosophie ebenfalls eine Rolle spielen, sollen hier nicht betrachtet werden.

Empfehlung 1:
Buchpublikationen (Monographien und Sammelbände) sollen in Zusammenarbeit mit Infrastruktureinrichtungen (i.d.R. Bibliotheken) an Universitäten oder Forschungseinrichtungen im Rahmen von Diamond-Open-Access-Modellen realisiert werden.

Um der Schwierigkeit eines anfänglichen Reputationsdefizits derartiger Publikationsmodelle entgegenzuwirken, wird vorgeschlagen, dass insbesondere die Fachgesellschaften selbst als ‚Zugpferde' tätig werden und geplante Veröffentlichungen aus Kongressen, Fachforen, Arbeitsgemeinschaften etc. über Diamond-OA-Modelle abwickeln.

Diamond-Open-Access meint dabei die Etablierung von Publikationsinfrastrukturen, bei denen Institutionen, Fachgesellschaften oder Förderorganisationen den dauerhaften Betrieb sicherstellen, sodass sowohl für AutorInnen als auch für LeserInnen keine Gebühren oder Bezahlschranken bestehen. Eine solche Infrastruktur kann zum einen komplett eigenständig mit (ggf. bereits bestehenden) Universitätsverlagen realisiert werden. Zum anderen kann durch Zusammenarbeit mit (kleinen) privatrechtlichen Verlagen in Co-Publishing-Ansätzen die Veröffentlichung wissenschaftlicher Ergebnisse auch ohne eigene technisch-organisatorische Verlagsstruktur durch öffentlich-rechtliche Institutionen ermöglicht werden.

Diese Publikationsweise erlaubt eine zeitnahe Umsetzung von Buchprojekten und die direkte, freie Zugänglichmachung von Forschungsergebnissen innerhalb der wissenschaftlichen Community, sodass nicht kommerzielle Verwertungsinteressen, sondern primär ein vereinfachter Austausch von Ideen im Vordergrund steht. Die Informationshoheit über erzielte Forschungsergebnisse würde damit zurück an die Einrichtungen

übertragen werden, an denen diese Ergebnisse auch erarbeitet worden sind.

Es bietet sich an, Services im Bereich des elektronischen Publizierens durch Angebote von Print-on-demand zu ergänzen, sodass auch weiterhin das gedruckte Buch als Austauschmedium zur Verfügung steht.

Berücksichtigt werden sollte bei der Umsetzung dieser Maßnahme, dass es hierdurch zu keiner Zusatzbelastung für WissenschaftlerInnen im Alltagsgeschäft kommen darf. Die Dienstleistungen des elektronischen Publizierens sollten vielmehr von Mitarbeitenden der Infrastruktureinrichtungen gesteuert werden. Dieser (neue) Aufgabenschwerpunkt würde ferner zur Stärkung der Rolle dieser Einrichtungen innerhalb der Hochschulen beitragen, und ihr Bestehen im digitalen Zeitalter sicherstellen.

In diese Modelle sollten bestehende unabhängige Wissenschaftsverlage (in Deutschland) nach Möglichkeit integriert werden. Ökonomische Verflechtungen, wie sie durch das Aufkaufen von Konkurrenten im Markt mittlerweile üblich sind, können aus der Perspektive der WissenschaftlerInnen nicht adäquat beurteilt werden. Vor diesem Hintergrund wäre es ein weiteres Aufgabenfeld für die Bibliotheken, einen genauen Überblick über die Verlagslandschaft zu bewahren.

Neben den genannten Buchpublikationen sollten in einem weiteren Schritt ebenfalls Zeitschriftenpublikationen als wichtiger Sektor des elektronischen Publizierens einer kritischen Betrachtung unterzogen werden. Im digitalen Zeitalter hat sich das Format einer Zeitschrift mit Platzlimitationen auf Grund von gedruckten Ausgaben überholt. Diese Feststellung trifft auf eine ganze Reihe weiterer Aspekte zu, die aus dem analogen ins digitale Zeitalter übertragen wurden, ohne deren Sinnhaftigkeit zu hinterfragen. Dazu zählen beispielsweise die Festlegung eines Seitenumfangs für einzelne Beiträge, der sich aus dem Platzmangel im gedruckten Medium ergab, oder vermeintliche Zusatzkosten für Farbabbildungen im Unterschied zu Graustufengrafiken, obwohl im elektronischen Dokument kein Kostenunterschied entsteht und anderes mehr.

2.2 Peer Review: Begutachtung nach Kriterien der Wissenschaftlichkeit

Das Peer-Review-Verfahren gilt gemeinhin als Gold-Standard zur Sicherstellung qualitativ hochwertiger Forschungsbeiträge. Wie in der Eingangsgrafik deutlich geworden ist, wird dieses Begutachtungsverfahren derzeit üblicherweise innerhalb der wissenschaftlichen Community als kostenlose Dienstleistung für Verlage etc. erbracht.

Das gegenwärtige Peer-Review-Verfahren weist dabei eine Reihe von Schwächen auf:

- Für die GutachterInnen ist es eine zeitaufwändige Zusatzaufgabe, die auf Grund der stetig steigenden Publikationszahlen zu einer wachsenden Belastung im Arbeitsalltag führt. Eine Folge davon ist, dass die Aufgabe an weniger qualifizierte GutachterInnen (wissenschaftlicher Nachwuchs) weitergereicht wird.
- Für die AutorInnen stellen Peer-Review-Prozesse oftmals langwierige Abfolgen immer neuer Überarbeitungsschritte dar. Dies trifft insbesondere auf Beiträge innerhalb der Geisteswissenschaften zu, da hier keine ‚hard facts' von den GutachterInnen kritisiert werden, sondern unterschiedliche Forschungsansätze und Paradigmen im Vorwege der Publikation diskutiert werden. Die Folge dieses defizitären Prozesses ist derzeit eine starke zeitliche Verzögerung, bis Arbeitsergebnisse der Community zugänglich gemacht werden können.
- Für AutorInnen, die sich noch in der Qualifikationsphase befinden, bedeutet dieser in die Länge gezogene Publikationsprozess eine zusätzliche Belastung, da sie sich mit ihren Arbeitsergebnissen auf dem akademischen Arbeitsmarkt bewerben müssen. Unnötig lange Feedback-Schleifen des Begutachtungsverfahrens stehen jedoch einer zeitnahen Publikation von erzielten Ergebnissen im Wege und begünstigen darüber hinaus eine übervorsichtige Themenwahl und Argumentationsweise, um KritikerInnen im Vorwege für die eigene Arbeit günstig zu stimmen.

Empfehlung 2:
Peer-Reviews zur qualitativen Bewertung von Zeitschriftenartikeln sollen nach expliziten (ggf. publikationsorganübergreifend standardisierten) Kriterien durchgeführt werden. Diese Kriterien zielen

primär auf den Nachweis der Wissenschaftlichkeit des Beitrags ab. Die Bewertung kann qualitativ in einem Formular mit gradueller Abstufung vorgenommen werden.

Darüber hinausreichende inhaltliche Vorschläge für eine Überarbeitung können im Begutachtungsverfahren auch weiterhin genannt werden (z.B. mittels einer Kommentarfunktion), sollten aber keinen Einfluss auf die Annahme oder Ablehnung eines zur Publikation eingereichten Artikels haben.

Kriterien, die ein Text in dieser Hinsicht erfüllen sollte, sind u.a.:

- Präzision der Darstellung und Klarheit der Begriffe
- Stringenz und Konsistenz der Argumentation
- Nachvollziehbarkeit der Darstellung (Plausibilität)
- Wissenschaftliche Relevanz und kritische Reflexion der Ergebnisse (Signifikanz)
- Theorieniveau und Lösungsqualität
- Methodenniveau: Wahl geeigneter philosophischer Methoden (z.B. Gedanken experimente, Begriffsanalyse usw.)
- Berücksichtigung alternativer Perspektiven und Umgang mit Gegenbeispielen

Die Erfassung der genannten Kriterien kann in einem Webformular vorgenommen werden, wie es bisher z.B. für die Feststellung der Relevanz eines Beitrags für ein bestimmtes Publikationsorgan schon erfolgt. Es ist ferner möglich, einige der Kriterien in gradueller Abstufung (‚vollständig erfüllt' bis ‚gar nicht erfüllt') angeben zu lassen. Die Arbeit mit einem solchen vorkonfigurierten Webformular könnte zu einer erheblichen Reduzierung des Zeitaufwands und zu einer Erhöhung der Transparenz für Publikationsannahmen bzw. -ablehnungen beitragen.

2.3 Bewerbungsverfahren: Nutzung von Visualisierungen statt Metriken

Im Rahmen von Bewerbungsgesprächen (Job-Interview, Berufungsverfahren etc.) sollte die Möglichkeit bestehen, den Kandidaten / die Kandidatin in angemessenem Umfang im Hinblick auf

die Passung für die ausgeschriebene Stelle zu evaluieren. Eindimensionale Zahlenwerte, wie sie durch bibliometrische Verfahren gewonnen werden (h-Index, Impact Faktor etc.), geben eine solche Würdigung der Diversität von beruflichen Lebensläufen nicht her. Darüber hinaus besteht bei diesen metrischen Werten das Problem, dass sie gezielt beeinflusst werden können („Salami-Taktik" der Publikationen, Teilnahme an Zitationskartellen etc.). Die Deutsche Forschungsgemeinschaft hat bereits mehrfach darauf hingewiesen, dass diese Zahlenwerte allenfalls zusätzlich in die Bewertung von KandidatInnen einbezogen werden, aber nicht ausschlaggebend für die Stellenbesetzung sein sollten.

Schließlich geht es auch darum, das Ineinandergreifen von elektronischem Publizieren und metrikbasierter Wissenschaftsevaluation zu lösen, um die Abhängigkeit von kommerziellen Datenbankbetreibern und Informationsanalysten zu beenden. Hierzu ist ein umfassender Kulturwandel der Bewertung von Fach-Expertise innerhalb der wissenschaftlichen Gemeinschaft selbst notwendig. Ein erster Schritt zur Änderung der derzeit vorherrschenden Präferierung numerischer Wertungen („Quantifizierung des Bewerbers / der Bewerberin") in der entsprechenden Kommissionsarbeit bestünde in einer qualitativen Erweiterung dieser Quantifizierungen.

Empfehlung 3:
In Bewerbungsverfahren zur Auswahl oder Berufung von KandidatInnen insbesondere durch Kommissionen sollen grafisch aufbereitete, mehrdimensional aussagekräftige Daten beispielsweise zu Themenbreite, Themenschwerpunkten und Publikationszeiträumen genutzt werden, anstatt lediglich eindimensionale Zahlenwerte zu vergleichen.

Diese mehrdimensionalen Daten können ebenfalls mittels Datenbankabfrage aus den Metadaten der Publikationseingaben eines Bewerbers / einer Bewerberin gewonnen werden. Die visuelle Darstellung von Arbeitsgebieten und der Publikationshistorie können sich dabei sehr einfach auf Datensammlungen in hochschuleigenen Repositorien oder community-basierten Datenbanken wie z.B. dem Preprint-Server „philpapers.org" stützen. Ein

solches Vorgehen fördert zudem eine Ablösung von kommerziellen Informationsanalysten und ihren IT-Produkten, da das im ersten Punkt der Empfehlungen angesprochene community-gestützte Publizieren im Rahmen von Diamond-OA-Modellen die Möglichkeit bietet, die Informationshoheit über die notwendigen Metadaten vollständig in die Hände der wissenschaftlichen Gemeinschaft zurück zu übertragen.

Im Kontext von Kommissionsgesprächen etc. könnten grafisch aufbereitete Zusammenstellungen von Metadaten zu Arbeitsgebieten, Publikationshistorie usw. als Leitfaden für Auswahlgespräche dienen. Anhand solcher Grafiken können eigene Arbeitsschwerpunkte, eine Spezialisierung von Forschungsinteressen oder eine Verlagerung derselben über die Zeit hinweg thematisiert werden, sodass Informationen über die fachliche Breite oder Spezialisierung eines Bewerbers / einer Bewerberin ausgetauscht werden können. Zusammenhänge zwischen den einzelnen Arbeitsbereichen können erörtert und in der Vita kenntlich gemacht werden. Die Verbindung zwischen z.B. Projektlaufzeiten und Publikationsergebnissen können angesprochen ebenso wie zeitliche Belastungen durch Lehre, Elternzeit etc. verdeutlicht werden. Auf diese Weise kann ein umfassenderes Bild der beruflichen Laufbahn des Bewerbers / der Bewerberin vermittelt und erfasst werden, als dies durch die Fokussierung vermeintlich objektiver Zahlenwerte aus bibliometrischen Daten möglich wäre.

Auf diese Weise kann auch vermieden werden, KandidatInnen allein auf Grund derartiger Zahlenwerte vorzeitig aus dem Bewerbungsverfahren auszusieben, ohne ihre tatsächliche inhaltliche Passung zu anderen Forschungsaktivitäten am Bewerbungsort in Betracht gezogen zu haben.

2.4 Promotionsphase: Trennung der Rollen von Betreuung und Begutachtung

Begutachtungsverfahren spielen darüber hinaus in der Ausbildung des wissenschaftlichen Nachwuchses eine wesentliche Rolle.

Empfehlung 4:
Um Abhängigkeiten und Machtmissbrauch im akademischen System vorzubeugen und abzubauen, sollen die Rollen von Betreuenden und Prüfenden in der Promotionsphase auf der personellen Ebene voneinander getrennt werden (vgl. britisches System). Sollten Beschäftigungsverhältnisse bestehen, dann wäre auch die Trennung von Fachvorgesetzten und Prüfenden wünschenswert.

Auf diese Weise kann die Neutralität der Begutachtung erhöht werden. Die Beurteilung im Gutachten zum Ergebnis (Text und Prüfung) erfolgt dabei unabhängig von den Höhen und Tiefen einer mehr oder weniger erfolgreichen Betreuungsphase.

2.5 Datenhoheit: Unterbindung des Wissenschaftler-Trackings

Der Ausbau kommerzieller Wissenschaftsverlage zu sogenannten Informationsanalysten führt dazu, dass diese immer wieder neue Marktsegmente erschließen, die durch Auswertung und Verkauf von NutzerInnen-Daten entstehen. Das „Wissenschaftler-Tracking" ist der jüngste Höhepunkt dieser Entwicklung. Die Auswertung des Informationssuch- und -nutzungsverhaltens von NutzerInnen auf den Plattformen der Anbieter liefert eine Menge von Daten, die in mehrerlei Hinsicht zum Nachteil der wissenschaftlichen Gemeinschaft und der individuellen WissenschaftlerInnen verwendet werden können.

Wie die Deutsche Forschungsgemeinschaft in ihrem Informationspapier „Datentracking in der Wissenschaft: Aggregation und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage" (2021) erörtert, besteht so beispielsweise die Gefahr, dass die Steuerung zukünftiger Forschungsaktivitäten nicht mehr direkt aus der Community selbst heraus erfolgt, sondern mit der Auswertung von Suchhistorien auf kommerziellen Datenbanken vermeintliche Forschungsziele abgeleitet werden. Gesucht und ausgewertet werden kann aber nur, was zuvor seinen Weg in diese Datenbanken schon gefunden hat, wodurch

etablierte Positionen strukturell bevorzugt werden. Aber nicht nur deshalb sollte die Zulässigkeit eines Induktionsschlusses von empirisch gemessenen Forschungsinteressen und -schwerpunkten auf Forschungsziele abgelehnt werden. Neben generellen Induktionsfehlschlüssen verbirgt sich hier auch ein Fehlschluss vom Sein aufs Sollen.

Für individuelle ForscherInnen ist die Auswertung ihrer Informationssuchen und –nutzungsweisen besonders problematisch, weil dadurch nicht nur ihre Forschungsergebnisse, sondern zusätzlich auch ihr Forschungsweg von anderen Akteuren beurteilt und bewertet werden kann. Besonders riskant für ForscherInnen ist das dann, wenn sich aus (gesellschafts-)politischen Gründen institutionelle oder staatliche Akteure gerade nicht der Forschungsfreiheit verpflichtet sehen.

Empfehlung 5:
Um Datentracking zu vermeiden und Datensouveränität zu etablieren, soll die Infrastruktur des elektronischen Publizierens und die damit verbundene Erhebung und Speicherung von Metadaten nicht nur rechtlich (Datenschutz) sondern auch in der institutionellen Trägerschaft (siehe Empfehlung 1) transformiert werden.

Die Realisierung dieses Vorhabens ist jedoch nur auf internationaler Ebene erfolgversprechend. Empfohlen wird vor diesem Hintergrund die Unterstützung von entsprechenden Initiativen wie der „Coalition for Advancing Research Assessment (CoARA)".

Anmerkungen

[1] Dieses Positionspapier wurde publiziert als Mößner, N., & Erlach, K. (2024). „Elektronisches Publizieren und Bewertung wissenschaftlicher Expertise". In: Zenodo. https://doi.org/10.5281/zenodo.10948472.

Literatur

Brembs, Björn; Philippe Huneman, Felix Schönbrodt, Gustav Nilsonne, Toma Susi, Renke Siems, Pandelis Perakakis, Varvara Trachana, Lai

Ma, & Sara Rodriguez-Cuadrado (2023) „Replacing academic journals“. *Zenodo*. https://doi.org/10.5281/zenodo.7974116.

Coalition for Advancing Research Assessment (CoARA): https://coara.eu/.

DFG-Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme (2021) „Datentracking in der Wissenschaft: Aggregation und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage. Ein Informationspapier des Ausschusses für Wissenschaftliche Bibliotheken und Informationssysteme der Deutschen Forschungsgemeinschaft“. *Zenodo*. https://doi.org/10.5281/zenodo.5900759.

Deutsche Forschungsgemeinschaft | AG Publikationswesen (2022) „Wissenschaftliches Publizieren als Grundlage und Gestaltungsfeld der Wissenschaftsbewertung“. *Zenodo*. https://doi.org/10.5281/zenodo.6538163.

DGPhil-Forumstagung „verTRACKTe Infrastruktur?! – Von Konstrukteuren und Kontrolleuren wissenschaftlicher Expertise“: https://www.izkt.uni-stuttgart.de/veranstaltungen/verTRACKte-Infrastruktur-Von-Konstrukteuren-und-Kontrolleuren-wissenschaftlicher-Expertise/.

Holzer, Angela (2022) „Die Vermessbarkeit der Wissenschaft. Digitalisierung, wissenschaftliches Publizieren, Verhaltenstracking und Wissenschaftsbewertung“. In *Kalibrierung der Wissenschaft. Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis,* hrsg. von N. Mößner und K. Erlach, Bielefeld: transcript, S. 163–181. https://doi.org/10.1515/9783839462102.

KAPITEL 13

Das Projekt Diamond_Thinking

Stefan Drößler

Im vorliegenden Band „Vermessene Philosophie“ werden nicht nur aktuelle Probleme des wissenschaftlichen Publizierens thematisiert, sondern es wird auch der Weg des *Diamond Open Access* (DOA) für die Veröffentlichung angewendet. Die Produktion wurde dementsprechend im Verbundprojekt Diamond_Thinking realisiert.[1] Dabei ging es u.a. darum, ein Publikationskonzept für Konferenzveröffentlichungen für die Fachkultur der Philosophie zu erproben, das alternativ zu kommerziellen Verlagen die Buchpublikationsservices der Bibliothek nutzt, aber auch einen externen Dienstleister einbezieht. Im Folgenden werden theoretische Grundlagen, Ziele und Konzepte des Projekts Diamond_Thinking vorgestellt.[2] Abschließend wird die konkrete Umsetzung bei der Produktion des Bandes zusammengefasst.

1. Wissenschaftstheoretischer Ansatz

Open Access als Paradigma zur Steigerung der Sichtbarkeit von Forschungsleistungen ist heute im Wissenschaftssystem

Zitationsvorschlag für dieses Kapitel:
Drößler, S. 2026. Das Projekt Diamond_Thinking. In: Mößner, N. und Erlach, K. (Hrsg.) *Vermessene Philosophie: Konstruktion und Kontrolle wissenschaftlicher Expertise im digitalen Raum.* S. 367–373. London: Ubiquity Press. DOI: https://doi.org/10.5334/bdh.p. License: CC BY 4.0

weitestgehend angekommen. Bibliotheken und kommerzielle Verlage ringen seit Jahren um die Ausgestaltung der Open-Access-Transformation. Allerdings bleibt die Finanzierung schwierig. Dabei steht bisher in erster Linie die Ausgestaltung neuer Geschäftsmodelle im Vordergrund, die auf die bestehenden oligopolartigen Strukturen und Preismodelle kaum nennenswerten Einfluss haben. Der Wissenschaftsrat und die Deutsche Forschungsgemeinschaft haben daher jüngst erneut Bibliotheken und Forschungseinrichtungen explizit aufgefordert, den Betrieb wissenschaftsgeleiteter Publikationsservices umfänglich zu unterstützen, um so einen Beitrag zu leisten, den teilweise gravierenden Verzerrungen auf dem Publikationsmarkt mit ihren negativen Auswirkungen auf die Wissenschaft entgegenzuwirken (vgl. Deutsche Forschungsgemeinschaft 2022; Wissenschaftsrat 2022). Damit stehen sie im Einklang mit vielen weiteren Aufrufen, das wissenschaftliche Publikations- und Kommunikationssystem inhaltlich neu auszurichten.

Die Alternativen werden unter dem Begriff „Diamond Open Access" (DOA) diskutiert. Darunter werden Finanzierungs- und Geschäftsmodelle für Publikationsinfrastrukturen gefasst, bei denen Institutionen, Fachgesellschaften und Förderorganisationen den Betrieb der Infrastrukturen sicherstellen, sodass für Autor:innen und Lesende keine Gebühren oder Bezahlschranken bestehen. DOA-Infrastrukturen sollten so angelegt sein, dass Publikations- und Kommunikationssysteme in der Hand der Wissenschaft liegen. Sie beinhalten Services und Standards rund um das interdisziplinäre und vernetzte wissenschaftliche Publizieren. Forschende profitieren von diesen Infrastrukturen durch qualitätsgesicherte und für sie kostenlose Publikationswege. DOA ist die einfachste und fairste Form von Open Access für alle Beteiligten, da der wissenschaftliche Diskurs von der Entstehung bis zur Verbreitung der Inhalte vollkommen ohne (Bezahl-)Schranken verläuft. Die Wissenschaft bestimmt in den eigenen Infrastrukturen selbst über Ausgestaltung der Servicelevels, Qualitätsstufen und Workflows.

2. Ziele und Umsetzung

Ziel des Projektes war es, die konkrete Ausgestaltung der Services für DOA-Alternativen zu entwickeln. Beide Projektpartner (KIT-Bibliothek und UB Stuttgart) gingen aufgrund der vielfachen Rückmeldungen von den Wissenschaftler:innen ihres Campus mit großer Sicherheit davon aus, dass die im vorliegenden Antrag projektierten DAO-Geschäftsmodelle bereits in den Köpfen der Community „schweben" und in vielen Fällen auf eine Anwendung regelrecht „warten".

An der Universität Stuttgart konnte an die Zusammenarbeit mit Nicola Mößner, Vertretungsprofessorin am Institut für Philosophie (für die Zeiträume April 2019 bis März 2020 sowie April 2022 bis September 2022) sowie Klaus Erlach vom Stuttgarter Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA angeknüpft werden. Im Rahmen des interdisziplinären Workshops „Wissenschaft im digitalen Raum – Erkenntnis in Filterblasen?"[3] wurden u.a. krisenhafte Entwicklungen im Bereich des wissenschaftlichen Publikationssystems diskutiert und Lösungsansätze in den Blick genommen (Mößner & Erlach 2022). Als ein nächster Schritt sollte die Zusammenarbeit zwischen der Universitätsbibliothek und der Fachdisziplin der Philosophie genutzt werden, alternative DOA-Publikationswege zu entwickeln.

Durch die bisherigen OA-Aktivitäten am KIT und der Universität Stuttgart hatten die Antragsteller von vielen Wissenschaftler:innen die Rückmeldung bekommen, dass sie gerne Open Access mit Hilfe ihrer Fachgesellschaften umsetzen würden, dazu aber viel Überzeugungsarbeit nötig sei, da die Vorstände oft national, wenn nicht sogar international besetzt sind. Die Diskussionen drehten sich entsprechend um eine neu zu definierende Open-Access-Kultur in der Wissenschaftsgemeinschaft.

Am Beispiel der Konferenzveröffentlichungen (Proceedings) wird deutlich, dass alte Strukturen wie die Publikation in Closed-Access-Zeitschriften oder das einfache „Veröffentlichen" über die Website nicht mehr gewünscht werden. Allerdings herrscht Uneinigkeit über das anzustrebende Publikationsmodell. In diesem

Zusammenhang müssen auch Bedenken hinsichtlich von Open-Access-Alternativen berücksichtigt werden. Typische Bedenken gibt es in Bezug auf die Kosten, die Reputation, die Qualitätssicherung und die Sichtbarkeit der Open-Access-Publikation, gerade wenn sie nicht in einem bekannten Wissenschaftsverlag erfolgt.

Die bisherigen Erfahrungen an beiden Standorten zeigen auch, dass die komplette Kette der OA-Dienstleistungen viele Brüche[4] aufweist, die auf beiden Seiten, sowohl bei der Wissenschaft als auch bei den Infrastrukturanbietern, viel Arbeit generieren. Zudem muss festgestellt werden, dass viel Know-how und Arbeitskraft an kommerzielle Verlagsanbieter fließen, wenn Wissenschaftler:innen als Reviewer tätig sind oder in Editorial Boards mitwirken.

Darüber hinaus muss festgestellt werden, dass die konkrete Umsetzung von Open Access bisher nicht nur an technischen Hürden scheitert, sondern auch an wissenschaftskulturellen Praxen und Einstellungen. Deshalb wurde die Erarbeitung der DOA-Transformation u.a. für die Deutsche Gesellschaft für Philosophie e.V. (DGPhil), die größte nationale Fachgesellschaft der Philosophie, in Form eines Anwendungsfalls evaluiert. Schließlich sollten Empfehlungen abgeleitet werden, wie die Erkenntnisse konkret nutzbar gemacht und zur Erreichung einer responsiven Open-Access-Umgebung eingesetzt werden können. Diese Empfehlungen werden zur Nachnutzung in anderen Einrichtungen veröffentlicht.

Konferenzveröffentlichungen können sowohl in Buchform als auch in Zeitschriften erscheinen. Vor diesem Hintergrund wurden für die Zeitschriftenoption sowohl der eigene Hostingservice für die Software Open Journal Systems mit den Wissenschaftler:innen geprüft als auch in Zusammenarbeit mit dem KIT (KIT Scientific Publishing, KSP)[5] und dem Dienstleister Ubiquity Press[6] das Co-Publishing-Modell für Buchpublikationen. Der Ansatz des Co-Publishing wurde bereits im Vorgänger-Projekt TU9_Monos (Drößler 2024) erprobt. Dabei hat sich die Veröffentlichung nach dem Baukastensystem besonders für kleinere Bibliotheken ohne Universitätsverlag als attraktiv erwiesen: Die Bedarfe können mit den Wissenschaftler:innen abgestimmt werden, sodass im

Unterschied zu einer feststehenden Book Processing Charge, wie in einem herkömmlichen Verlag, nur Services in Anspruch genommen werden müssen, die auch wirklich gebraucht werden. Hierdurch wurde es möglich, gebührenfreie DOA-Alternativen für die Forschenden anzubieten und Leistungen, die die eigenen Basisdienste übersteigen, zum Beispiel Indexierungsservices, extern einzukaufen oder kooperativ im Verbund mit anderen Universitäten zu lösen.

3. Produktion des Bandes

Für die Produktion des Sammelbandes „Vermessene Philosophie" war zunächst geplant, Basisdienste wie Satz und Layout mit den im Projekt Diamond_Thinking zur Verfügung stehenden Ressourcen zu leisten. Dafür wurde eine Manuskriptvorlage erstellt. Durch längere Abstimmungsprozesse mit Ubiquity für die Vertragsgestaltung und das Review-Verfahren wurde aus Zeitgründen entschieden, nicht nur Indexierungsservices (Aufnahme in Kataloge und Datenbanken wie DOAB[7] und JSTOR[8]) und Veröffentlichung auf der E-Publishing-Plattform von Ubiquity Press machen zu lassen, sondern auch Satz und Layout.

Die Umschlaggestaltung für das Buch hat KIT Scientific Publishing übernommen in Form einer Kulanzregelung gegenseitiger Unterstützung innerhalb der Projektzusammenarbeit. Dafür wurden zwei Entwürfe mit thematischem Bezug zur Grafik mit den Herausgebern abgestimmt.

Die Herausgeber übernahmen neben der Konzeption des Bandes die Kommunikation mit den Autor:innen, warben zusätzliche Beiträge ein und begutachteten die eingereichten Manuskripte. Das für die Fachkultur der Philosophie übliche Qualitätssicherungsverfahren des Herausgeber-Reviews wurde Ubiquity Press transparent gemacht, um den Standards des Verlags gerecht zu werden.

Normalerweise ist es bei Publikationsprojekten notwendig, die geeignete Open-Access-Lizenz zu diskutieren und ein gemeinsames Verständnis zu entwickeln. Durch die intensive Beschäftigung

mit dem Open-Access-Publizieren sowohl von Seiten der Herausgeber als auch auf Seiten der Bibliothek und des Open-Access-Verlags Ubiquity Press, war an dieser Stelle ein Konsens schnell gefunden. Die Finanzierung konnte nach Rücksprache mit dem Projektträger VDI/VDE aus Projektmitteln erfolgen. Die Kosten werden nach dem Erscheinen des Bandes veröffentlicht.

Anmerkungen

[1] Das Verbundprojekt Diamond_Thinking wurde vom Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) im Zeitraum von September 2023 bis August 2025 gefördert. Verbundpartner waren die Bibliotheken des Karlsruher Instituts für Technologie (KIT) und der Universität Stuttgart. Für das Teilvorhaben Stuttgart wurde die Laufzeit bis Ende 2025 verlängert.

[2] Der Projektantrag, auf den sich die nachfolgende Ausarbeitung stützt, wurde von Regine Tobias (ORCID: https://orcid.org/0000-0002-6164-0832), Sibylle Hermann (ORCID: https://orcid.org/0000-0001-9239-8789) und Stefan Drößler formuliert. Der vorliegende Beitrag enthält bisher nicht-veröffentlichte, gekürzte und moderat angepasste Teile des Antrags.

[3] Vgl. zum Programm: https://www.philo.uni-stuttgart.de/digitalerraum.html (eingesehen am 12.01.2026).

[4] Brüche gibt es zum Beispiel im Workflow der Manuskriptbearbeitung, wenn Autor:innen, Gutachter:innen und Grafiker:innen mit unterschiedlicher Software arbeiten. Bei der Bearbeitung von Satz und Layout verwenden Grafiker:innen meist Desktop-Publishing-Software (DTP), d.h. die mit Standard-Textbearbeitungssoftware (Libre Office, Word) geschriebenen Manuskripte müssen mit DTP von Grund auf neu gesetzt werden, was arbeitsintensiv ist. LaTeX-Manuskripte können für die Weiterbearbeitung besonders dann problematisch werden, wenn die Vorlagen nicht hinreichend dokumentiert sind und beim Kompilieren Fehler auftauchen, die behoben werden müssen. So gibt es Abhängigkeiten zwischen der installierten TeX-Live-Version und den verwendeten TeX-Paketen und dem Compiler. Häufig schreiben Wissenschaftler:innen eigene TeX-Klassen und Pakete, die dann auch wieder Abhängigkeiten aufweisen.

[5] Vgl. https://www.ksp.kit.edu.

[6] Vgl. https://ubiquity.pub.
[7] Vgl. https://www.doabooks.org.
[8] Vgl. https://www.jstor.org.

Literatur

Deutsche Forschungsgemeinschaft, AG Publikationswesen (2022) Wissenschaftliches Publizieren als Grundlage und Gestaltungsfeld der Wissenschaftsbewertung. https://doi.org/10.5281/ZENODO.6538163.

Drößler, S. (2024) BMBF-Projekt TU9_Monos – Aufbau von nachhaltigen Open-Access-Publikationsservices für Monographien im TU9-Verbund, Teilvorhaben Universität Stuttgart: Schlussbericht. Universität Stuttgart, Universitätsbibliothek. https://doi.org/10.2314/KXP:1917824394.

Mößner, N.; Erlach, K. (2022) Kalibrierung der Wissenschaft: Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis. Transcript. https://doi.org/10.14361/9783839462102.

Wissenschaftsrat (2022) Empfehlungen zur Transformation des wissenschaftlichen Publizierens zu Open Access. https://doi.org/10.57674/FYRC-VB61.

Verzeichnis der AutorInnen

Andersen, Hanne (Prof. Dr.) ist Professorin für „history and philosophy of science" an der Universität von Kopenhagen. Ihr hauptsächlicher Forschungsschwerpunkt liegt auf historischen Entwicklungen und strukturellen Bedingungen der Wissenschaften im 20. und 21. Jh. (im weitesten Sinne verstanden). In ihren Untersuchungen der wissenschaftlichen Praxis verbindet sie historische und gegenwärtige Fallstudien mit analytischen Werkzeugen der Philosophie und mit empirischen, methodologischen und theoretischen Einsichten aus den Bereichen der Wissenschaftssoziologie, Forschungen aus dem höheren Bildungswesen sowie der Bibliotheks- und Informationswissenschaften. ORCID: https://orcid.org/0000-0003-2050-4540

Drößler, Stefan (M.A.) studierte Politikwissenschaft, Soziologie und Fachübersetzen Französisch. Zuletzt war er Fachreferent für Sozialwissenschaften und Philosophie an der Universitätsbibliothek Stuttgart sowie Open-Access-Beauftragter (2015–2025). Zu seinen Aufgaben gehörte die Weiterentwicklung der

Open-Access-Publikationsdienste über Drittmittelprojekte (DFG und BMBF). Seit Juli 2025 ist er stellvertretender Direktor an der Universitätsbibliothek Tübingen. ORCID: https://orcid.org/0000-0001-8071-5070

Erlach, Klaus (Dr.) ist organisatorisch und methodologisch verantwortlich für das Doktorandenprogramm am Fraunhofer Institut für Produktionstechnik und Automatisierung IPA in Stuttgart. Am Institut für Industrielle Fertigung und Fabrikbetrieb (IFF) der Universität Stuttgart leitet er die Gruppe „Theorie der Produktion". Zudem ist er Lehrbeauftragter an der Hochschule Vorarlberg und Redakteur des Philosophie-Journals „der blaue reiter". Seine Forschungsschwerpunkte umfassen die Wissenschaftstheorie der Technologie, Technikphilosophie sowie Gestaltungsmethoden zur Fabrikplanung und Produktionsoptimierung mit Wertstromdesign. Sein Buch „Das Technotop – Die technologische Konstruktion der Wirklichkeit" (2000) befasst sich mit dem technisch geprägten Lebensort des Menschen. Ein Standardwerk in Lehre und industrieller Anwendung ist sein Buch „Wertstromdesign – Der Weg zur schlanken Fabrik" (2007/^{2}2010/^{engl}2013 /^{korean}2018/^{3}2020). ORCID: https://orcid.org/0000-0002-9976-7468

Hangel, Nora (Dr.) vertritt (2026–27) die Professur für „Soziale Erkenntnistheorie und Argumentation in den Wissenschaften" an der RU Bochum. Sie ist PI des Projekts „Die Rolle wissenschaftlichen Beurteilung in der Wissenserzeugung: eine qualitative Studie über interpersonelle und kollektive Urteilsbildung in der wissenschaftlichen Praxis" (2022–25) am Leibniz Forschungszentrum Wissenschaft und Gesellschaft (LCSS) der Leibniz Universität Hannover. Mit einer Arbeit über I. Kant promovierte sie in Philosophie an der Universität Wien und war (2025) Gastwissenschaftlerin am Center for Philosophy of Science der Universität Pittsburgh. Artikel erschienen u.a. in: Studies in History and Philosophy of Science, EJPS, SSM- Qualitative Research in Health, Heliyon ORCID: https://orcid.org/0000-0001-6809-4603

Kaminski, Andreas (Prof. Dr.) ist Professor für Wissenschafts- und Technikphilosophie an der TU Darmstadt, er war Gastprofessor an der RWTH Aachen und ist Senior Scientist am Bundeshöchstleistungsrechenzentrum der Universität Stuttgart (HLRS), wo er zuvor eine Abteilung für Philosophie der computerbasierten Wissenschaften aufbaute. Zu seinen Forschungsgebieten zählen in der sozialen Epistemologie die Philosophie von Vertrauen und Zeugenschaft, in der technischen Epistemologie die Rolle der Technik in wissenschaftlichen Prozessen (Computersimulation und maschinellem Lernen) sowie ferner der Einsatz von Modellen für die Entscheidungsfindung (Modeling for Policy). Er ist Mitherausgeber des Jahrbuch Technikphilosophie. ORCID: https://orcid.org/0000-0002-9338-2575

Mittermaier, Bernhard (Dr.) ist promovierter Chemiker und hält einen M.A. in Library and Information Sciences. Er leitet die Zentralbibliothek des Forschungszentrums Jülich. Arbeitsschwerpunkte sind Open Access, Lizenzverhandlungen (auch im Rahmen von DEAL) und Bibliometrie; zu diesen Themenbereichen sind auch zahlreiche Publikationen erschienen https://juser.fz-juelich.de/search?ln=de&p=aid:P%3A(DE-Juel1)133810. ORCID: https://orcid.org/0000-0002-3412-6168

Mößner, Nicola (PD Dr.) ist Vertretungsprofessorin für „Theoretische Philosophie" am Institut für Philosophie der Leibniz Universität Hannover und Privatdozentin an der RWTH Aachen. Ihre Forschungsschwerpunkte umfassen die Wissenschaftsphilosophie, die soziale Epistemologie sowie die analytische Bildtheorie. Zu ihren wichtigsten Publikationen zählen „Wissen aus dem Zeugnis anderer – der Sonderfall medialer Berichterstattung" (2010), „Knowledge, Democracy, and the Internet" (2017, zus. mit Philip Kitcher), „Visual Representations in Science – Concept and Epistemology" (2018), „Kalibrierung der Wissenschaft. Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis" (2022, zus. mit Klaus Erlach). ORCID: https://orcid.org/0000-0002-3041-3071

Retzlaff, Eric (Dipl.-Bibl. (FH), M.A.) studierte Informations- und Medienmanagement. Er leitet am Fraunhofer IRB Stuttgart das Transformation Innovation Center (TIC) und ist stellvertretende Ombudsperson am Institut. Er arbeitete jahrelang im Informations- und Wissensmanagement und hält zahlreiche Seminare zu Unterstützungsdienstleistungen in Wissenschaft und Forschung. Als Dozent an der Hochschule der Medien (HdM) in Stuttgart für den Studiengang Informationswissenschaften vertritt er dort die Themen „Open Science & Open Access" und „Forschungsnahe Dienstleitungen". ORCID: https://orcid.org/0000-0002-5133-9544

Stein, Lisa-Marie (M.A.) studierte Linguistik, Deutsch und Medien mit dem Fokus auf Filmwissenschaft an der Christian-Albrechts-Universität zu Kiel. Seit 2020 ist sie im Bereich des wissenschaftlichen Publizierens tätig und aktuell als wissenschaftliche Mitarbeiterin in dem Projekt „openCost" in der Bibliothek des Forschungszentrums DESY in Hamburg beschäftigt. Berufsbegleitend vertieft sie die bibliothekswissenschaftliche Perspektive mit einem Studium der Informationswissenschaft (M.A. LIS) an der Humboldt-Universität zu Berlin. Ihre Interessen liegen insbesondere auf fairer Wissenschaftskultur, Open Science, Sprachen und Multilingualität. ORCID: https://orcid.org/0000-0001-7905-0462

Wiengarn, Jörn (Dr.) ist wissenschaftlicher Mitarbeiter am Institut für Philosophie der TU Darmstadt. Nach seiner Promotion an der Universität zu Köln war er als wissenschaftlicher Mitarbeiter im Projekt „Vertrauen in Informationen" am Bundeshöchstleistungsrechenzentrum der Universität Stuttgart (HLRS) tätig. Seine Forschungsinteressen liegen im Bereich der Sozialen Epistemologie sowie Wissenschafts- und Technikphilosophie. Derzeit beschäftigt er sich insbesondere mit Fragen des öffentlichen Vertrauens und Misstrauens in die Wissenschaft sowie mit Fragen zur Mis- und Desinformation. Er ist Redakteur des Jahrbuch Technikphilosophie. ORCID: https://orcid.org/0000-0002-6580-6019

STUTTGART UNIVERSITY PROCEEDINGS

VERMESSENE PHILOSOPHIE

In den letzten Jahrzehnten hat sich das wissenschaftliche Publizieren auch in der Philosophie vom gedruckten Werk hin zu überwiegend digitalen Formaten entwickelt.

Warum aber sollte man sich in der Philosophie mit Fragen zur Infrastruktur an Universitäten befassen? Wieso sollte man als Vertreter:in dieses Faches kritisch über Prozesse des elektronischen Publizierens nachdenken? Was haben Philosoph:innen mit Bibliometrie zu tun und wie verändert das die philosophische Arbeit? Die Autor:innen des hier vorgelegten Bandes diskutieren diese Fragen kritisch und entwickeln Vorschläge für mehr Autonomie in den Wissenschaften.

ISBN 978-1-914481-78-9
90000
9 781914 481789

Gedruckt auf FSC-zertifiziertem Papier

]u[ubiquity press