

Auf der Suche nach der verlorenen Qualität

Im Herbst will der Wissenschaftsrat Peer-Review-Empfehlungen geben. Im Vorfeld bilanziert die duz den Stand der Diskussion und fragt: Was ist zu tun, damit die Selbstreinigungskräfte der Wissenschaft wieder die gute wissenschaftliche Publikationspraxis sicherstellen können? Antworten der Vorsitzenden des Wissenschaftsrats, von Wissenschaftlern, Wissenschaftsmanagern und Experten lesen Sie in diesem Schwerpunkt.

von Pascale Anja Dannenberg



Foto: Gina Sanders/Fotoila

Es wird viel zu viel publiziert: Immer mehr Forschungsmanuskripte werden bei Fachzeitschriften eingereicht, sodass immer mehr Gutachten erstellt und immer mehr Ablehnungen ausgesprochen werden müssen – was zu noch mehr Anträgen führt. Die von deutschen und österreichischen Bibliotheken erstellte Zeitschriftendatenbank ZDB listet 1,8 Millionen Titel fortlaufender Sammelwerke, allen voran Zeitschriften und Zeitungen. Alle 20 Sekunden erscheinen ein Forschungsaufsatz und jeden Tag mehr als 240 Bücher, hat die Frankfurter Allgemeine Zeitung 2013 errechnet. Allerdings ziehen Fachzeitschriften 15-mal häufiger als ein Jahrzehnt zuvor fragwürdige Artikel zurück, das war ebenfalls 2013 zu lesen, in der Zeit. Und nur die Hälfte der in zitierten Zeitschriften erschienenen Artikel würden von anderen als den Autoren und den Gut-

„Die kritischen Stimmen nehmen zu“

achtern gelesen, und 90 Prozent der Artikel niemals zitiert, zählte der Wirtschaftswissenschaftler Prof. Dr. Alfred Kieser 2016.

„Früher reichte ein Wissenschaftler einen Aufsatz zur Veröffentlichung ein, wenn er der Meinung war, es hätten sich Ergebnisse seiner Forschungen angesammelt, die wichtig genug seien, um sie den Kollegen mitzuteilen. Heute reicht er einen Aufsatz ein, wenn er meint, er hätte ausreichend Stoff für eine Veröffentlichung beisammen, wobei er bemüht ist, aus seinen Ideen möglichst viele Aufsätze zu generieren“, beschreibt Kieser den heutigen Publikationsdruck. Das New Public Management zwingt Universitäten und Wissenschaftler in einen Verdrängungswettbewerb mit Rankings als Währungen. Er beobachtet, dass „die kritischen Stimmen an Zahl und Heftigkeit zunehmen“ und hofft auf ein „Umsteuern“.

Die Quantität gefährdet die Qualität. Autoren werden des wissenschaftlichen Fehlverhaltens überführt, Auszeichnungen müssen aberkannt werden, Studien stellen sich

als nicht replizierbar heraus. Verleger ziehen, möglichst still und leise, gedruckte Aufsätze zurück. Wissenschaftler, die durch ihr Gutachten den Daumen für die Annahme des Papiers eines Kollegen heben oder senken (Peer Review) und damit entscheidenden Einfluss ausüben, sind überlastet. Oftmals erhalten sie mehrere Anfragen täglich. Manche Wissenschaftler prangern den Publikationsdruck an, weisen auf Mängel beim Peer Review hin, indem sie als Autoren bewusst gefälschte Studien einreichen, um Herausgeber und Gutachter zu narren und öffentlich vorzuführen. Oder aber sie rufen eigene Online-Journals ins Leben, mit denen transparentere Begutachtungsprozesse einhergehen sollen. Zugleich wittern Pseudoverleger unseriöser Onlinemagazine der auf Beall's List aufgeführten predatory journals die profitable Gunst der Stunde und veröffentlichen mittels eines „alibihaften Peer-Review-Verfahrens praktisch alle Einreichungen“, sagt der Wiener Wirtschaftswissenschaftler Prof. Dr. Ulrich Berger. Nach seiner Einschätzung „führt das zu einer Publikationsschattenwelt zigtausender unseriöser, falscher und grotteschlechter Fachartikel, die nur existieren, weil immer noch in vielen akademischen Institutionen die reine ‚Anzahl der in Zeitschriften mit Peer-Review-Verfahren veröffentlichten Artikel‘ ein formales Kriterium für Beförderungen ist“.

Die Zunahme all dieser Krisensymptome macht deutlich, hier steht nicht mehr nur die Reputation einzelner Autoren, Herausgeber und Gutachter auf dem Spiel, sondern das Ansehen des gesamten Wissenschaftssystems. Was ist der Grund? Als Ersatz für am Markt erprobte Produkte der freien Wirtschaft bewertet die Wissenschaft sich selbst, sie legt durch Peer Reviews fest, was gute Forschung ist, analysiert die Ökonomin Prof. Dr. Margit Osterloh. Diese Gutachten werden herangezogen bei Entscheidungen über Zeitschriften- und Buchveröffentlichungen, auch bei Stellenbesetzungen und der Vergabe von Forschungsmitteln. Dabei gilt das Double-Blind-Peer-Review (s. S. 26 Kasten) geradezu als „heilige Kuh“, sagen Osterloh und Kieser, da dieses anonymisierte Bewertungsverfahren für sogenann-

Weiterlesen

Alfred Kieser: Ineffizient, irreführend und teuer: wissenschaftliche Zeitschriften in der Krise. Die Betriebswirtschaft DBW, 76. Jahrgang, 6/2016, S. 467ff.

Margit Osterloh: Würfelt Gott? Würfelt die Wissenschaft? Beiträge zur Hochschulforschung, 39. Jahrgang, 1/2017, S. 30ff.

Bruno S. Frey: Gutachten im Wissenschaftsprozess. Soziologie, Jahrgang 34, 2/2005, S. 166ff.

Bruno S. Frey, Margit Osterloh: Würfeln in der Wissenschaft? Forschung und Lehre, 23. Jahrgang, 2/2016, S. 134ff.

Margit Osterloh, Bruno S. Frey: Rankings und der Preis der Wissenschaft. Zeitschrift für Kulturwissenschaften, 1/2015, S. 65ff.

Margit Osterloh, Alfred Kieser: Double-Blind Peer Review: How to Slaughter a Sacred Cow. In: I. Weppe, J. Wollersheim, S. Ringelhan, M. Osterloh: Incentives and Performance: Governance of Research Organizations. Heidelberg, Springer 2015, S. 307ff.

Nikolaus Kriegeskorte: Open evaluation: a vision for entirely transparent post-publication peer review and rating for science. 2012.



Download

<https://tinyurl.com/yc3mm8yr>

te A-Journals gilt, die mit einem hohen Impact Factor gelistet sind – und damit der Maßeinheit, die angibt, wie oft im Durchschnitt alle Artikel dieser einen Zeitschrift im Zeitraum von zwei Jahren nach ihrer Veröffentlichung zitiert werden. Schon viele Jahre wird kritisiert, dass aus dem Impact Factor einer Zeitschrift ein Rückschluss auf die Qualität eines einzelnen Artikels gezogen wird. Manche Aufsätze werden oft zitiert, andere selten oder nie. Oder aber es spricht nach kurzem, oft zufälligem Hype niemand mehr von einer anfänglich häufig genannten Studie. Und so mancher Artikel, der große Schwierigkeiten hatte, den Gutachterprozess zu passieren, wurde später als Meilenstein anerkannt und führte gar

schen 12 000 Wissenschaftler und 851 Wissenschaftsorganisationen dafür eintreten, die Qualität eines Aufsatzes nicht nach dem Impact Factor zu bewerten. Und dennoch sind bis heute dieser und andere zahlenorientierte Leistungsindikatoren weiterhin die Karrierebeschleuniger auf dem Weg zur Habilitation oder Professur. So bequem solche Kennzahlen für Hochschulen, Forschungsfördereinrichtungen und Drittmittelgeber sein mögen, sie sind es letztlich, die zweifelhafter sind als die Peer Reviews selbst, da sie „die für den wissenschaftlichen Diskurs dringend erforderliche Vielfalt durch Einfachheit ersetzen“, sagt Margit Osterloh.

Was also ist zu tun, damit die Selbstreinigungskräfte der Wissenschaft wieder für Qualität bürgen können? Schon viele Reformvorschläge sind gemacht worden, manche wurden verworfen, neue werden gemacht. Gerade in Krisenzeiten entstehe oft vorschnell eine Begeisterung für Alternativen, beobachten die Professoren Dr. Stefan Hornbostel und Dr. Martin Reinhart

vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung. Doch stützen die Social-Media-Plattformen in keinem „sinnvollen Verhältnis von Qualität und Aufmerksamkeit“. Denn „wer den fragilen und sensiblen Prozess des bestehenden Peer Review ersetzen oder ergänzen will, tut gut daran, dieses nicht als ein Verfahren misszuverstehen, das Qualität zweifelsfrei erkennt. Die Leistung des Peer Review besteht vielmehr darin, Qualität zu produzieren und zu legitimieren.“ Im Peer Review werde ein geschützter Raum erzeugt, in dem ein kompetenter und kritischer Dis-

kurs entstehen solle. „Hier liegt aber auch die Zwickmühle: Ein geschützter Raum geht immer einher mit Intransparenz und Anonymität. Hohe und anerkannte Qualität der Begutachtung verlangt aber umgekehrt nach Offenheit, Nachvollziehbarkeit, Transparenz und gegebenenfalls nach Revidierbarkeit. Beides in ein ausgewogenes Verhältnis zu bringen, ist das eigentliche Kunststück.“ Dieser Kunst sei es allerdings keineswegs abträglich, wenn in strittigen oder konkurrierenden Fällen zu ungewöhnlichen Verfahren wie einer Lotterie gegriffen werde. „Wenn deutlich kommuniziert wird, dass trotz solider Begutachtung ein Entscheid nicht möglich ist, schädigt ein Losverfahren das Ansehen des Peer Review nicht, im Gegenteil, es entlastet von überzogenen Erwartungen.“

Kontrollierter Zufall

Der Vorschlag rekurriert auch auf Überlegungen von Margit Osterloh und Bruno S. Frey. Sie plädieren für eine partiell zufällige Auswahl von Personen oder Forschungsprogrammen. Dies sei die einzige Alternative, die, wenn die Zufallsauswahl kontrolliert durchgeführt werde, Favoritismus und Manipulation vermeide, die Diversität von Ideen gewährleiste und eine „Suchmaschine“ für neue Perspektiven und Talente eröffne, die im herkömmlichen Betrieb wenig Chancen hätten. „Zufall“ wollen die beiden Ökonomen hier im Sinne einer statistischen Wahrscheinlichkeit mit strengen mathematischen Gesetzmäßigkeiten verstanden wissen, fern jeder Willkür. Da der klare Nachteil des Verfahrens darin liege, dass nicht zwischen fähig und unfähig unterschied-

„Akademische Prostitution“

zum Nobelpreis. „Der Impact Factor ist ein grottenschlechter Qualitätsindikator“, bilanziert deshalb Kieser; und nach Ansicht des Schweizer Wirtschaftswissenschaftlers Prof. Dr. Bruno S. Frey zwingt er Wissenschaftler „zum Verkauf ihrer Seele“, schon 2005 spricht Frey daher auch von „akademischer Prostitution“.

Für die Vielfalt

So ist denn auch die Liste der Unterzeichner der 2012 initiierten Dora-Deklaration (San Francisco Declaration of Research Assessment) weiter angewachsen, nach der inzwi-

Volkswagen-Stiftung erprobt Losverfahren

■ Während einer vierjährigen Erprobungsphase entscheidet nicht allein eine Jury über Förderanträge, vielmehr wird der Zufall einbezogen: „Experiment - Auf der Suche nach gewagten Forschungsideen“ fördert 18 Monate lang Natur-, Ingenieur- und Lebenswissenschaftler mit 120 000 Euro. Bis 2020 sollen nicht nur aus einer Shortlist 15 bis 20 „überzeugende unkonventionelle Hypothesen“ per Double-Blind Peer Review ausgewählt werden, sondern zusätzlich aus dersel-

ben Shortlist in derselben Anzahl Anträge per Los. Wenn die geförderten Projekte abgeschlossen sind, sollen die Verfahren vergleichend evaluiert werden.

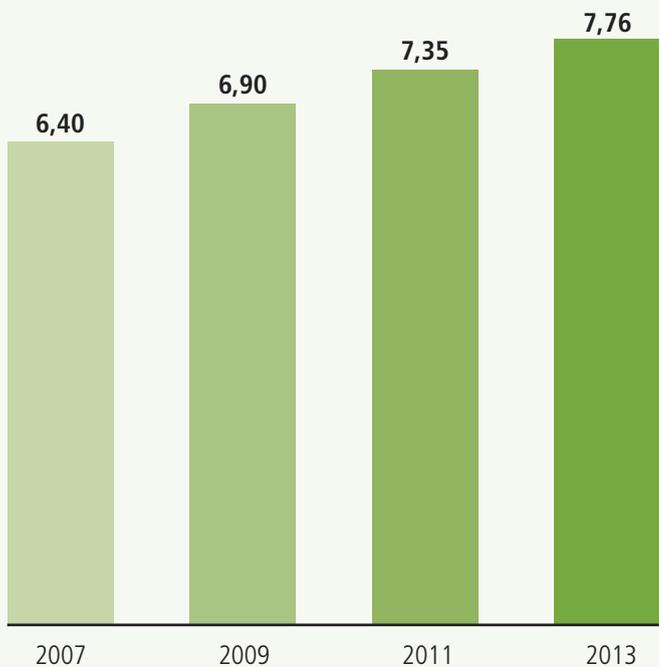
Stichtag 2017 ist der 5. Juli.

Internet

www.volkswagenstiftung.de/experiment.html

Zahl der Forscher weltweit

Angaben in Millionen



Quelle: Unesco, tinyurl.com/y9wa7m78

Die Zahl der Publikationen, die in der Literaturdatenbank Web of Science gelistet sind, wuchs **zwischen 2008 und 2014** um

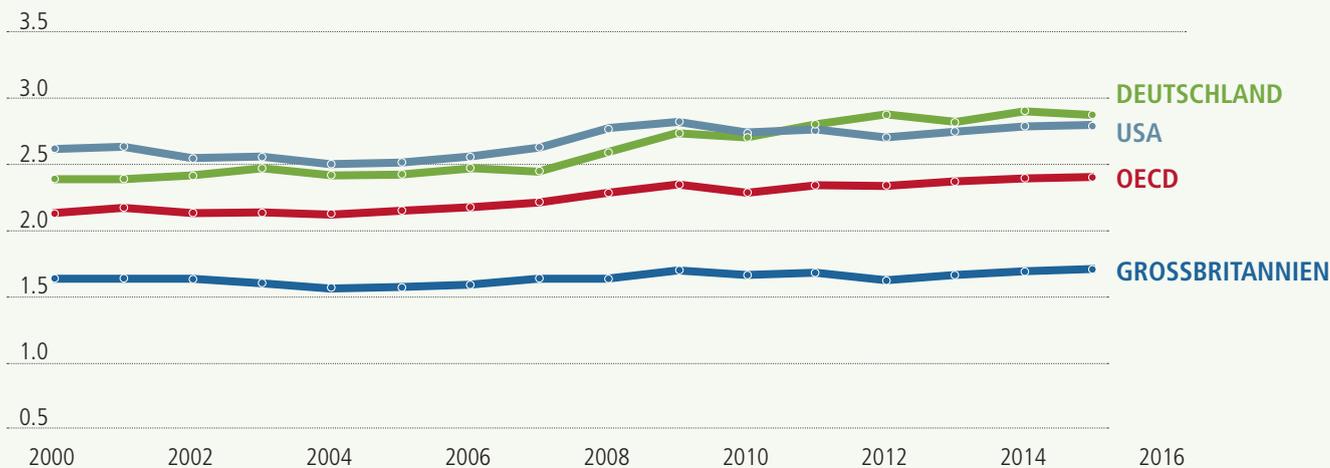
23 %

von **1 029 471** auf **1 270 425**.

Die **EU-Staaten** haben im Jahr 2014 432 195 wissenschaftliche Fachartikel veröffentlicht, das sind 34 % aller Artikel weltweit.

Quelle: Unesco, en.unesco.org/node/252282

Forschungs- und Entwicklungsausgaben ausgewählter Länder gemessen am BIP



Grafik: ESM/Quelle: OECD, tinyurl.com/o4vh567

Peer-Review-Modelle

Einseitige Blind-Begutachtung

Beim Single-Blind-Peer-Review ist dem Gutachter die Identität des Autors bekannt, jedoch nicht vice versa.

- + Einfach – und wohl auch deshalb das am häufigsten verbreitete Modell.
- Gefahr der Voreingenommenheit; Gutachter sind für Autoren über die (Auszüge aus) Gutachten zu erahnen.

Doppelt-blind-Begutachtung

Beim Double-Blind-Peer-Review kennt weder der Gutachter den Autor noch umgekehrt der Autor den Gutachter.

- + Die Gefahr der Voreingenommenheit ist geringer als beim Single-Blind-Peer-Review.
- Autoren sind für Gutachter zu erahnen (über Eigenzitationen); solche Hinweise effektiv zu verschleiern ist für Zeitschriften mit Kosten und Aufwand verbunden.

Offenes Peer Review

Ein Überbegriff für neuere Verfahren, die auf mehr Transparenz zielen. Teils werden Kommunikationsprozesse offengelegt, teils Gutachten, Datensätze, Revisionen und/oder auch die Namen der Autoren und Gutachter.

- + Befürworter glauben, dass alle Beteiligten gewissenhafter agieren, wenn ihre Kommunikation unter den Augen der Öffentlichkeit stattfindet.
- Manche Gutachter fürchten, angefeindet zu werden, wenn sie öffentlich Kritik üben.

Peer Review nach Veröffentlichung

Das Post-Publication-Peer-Review ist eine besondere Spielart des offenen Peer Review: Es erlaubt zusätzliche Gutachten und Kommentare von allen interessierten Lesern sowie Überarbeitungen des Autors nach der Veröffentlichung (s. S. 30-31).

- + Kontroversen werden öffentlich ausgetragen.
- Das Verfahren setzt auf Beteiligung, fordert damit Mehraufwand ein.

Vorangemeldete Studien

Beim Pre-Publication-Peer-Review kündigt der Autor einem Journal eine Studie an. Diese wird veröffentlicht, selbst bei einem negativen Ergebnis.

- + Zielt darauf, nicht allein positive und aufsehenerregende Ergebnisse zu veröffentlichen, sondern auch Rückschläge, denen ein ebenfalls hoher Erkenntniswert zugesprochen wird.
- Kritiker halten das Verfahren für einige Anwendungsbereiche geeignet (klinische Studien), häufig aber nicht für Grundlagenforschung.

Partielle Zufallsauswahl

In strittigen Fällen entscheidet ein Losverfahren (s. S. 24 u. 26)

- + Das Verfahren will kreativen, aber kontroversen Ansätzen eine Chance einräumen und eine Vielfalt in der Forschung stärken.
- Der Prozess unterscheidet nicht zwischen fähigen und unfähigen Forschern oder Forschungsanträgen.

hif

den werde, sollten zunächst mittels einstimmiger kurzer Gutachtereinschätzungen schlechte und gute Kandidierende beziehungsweise Anträge auf Forschungsmittel abgelehnt oder angenommen werden. „Die Zufallsauswahl könnte dann in den (meist überwiegenden) Fällen zur Anwendung kommen, bei denen Dissens herrscht. Es ist zu vermuten, dass es sich hierbei nicht selten um neuartige und ungewöhnliche Personen oder Beiträge handelt, die ansonsten wenig Chancen haben, sich im etablierten Wissenschaftsbetrieb durchzusetzen“, meint Osterloh. Und Frey ergänzt: „Nach fünf oder mehr Jahren könnten dann die zufällig ausgewählten Beiträge mit den nach dem üblichen Prozess ausgewählten Beiträge anhand von Zitierungen verglichen werden. Ich vermute, dass kein Unterschied feststellbar ist oder dass sogar die zufällig Ausgewählten öfter zitiert werden.“

Mehr Grundfinanzierung

Die Deutsche Forschungsgemeinschaft (DFG) appellierte 2015 an die Politik, die Grundfinanzierung an Hochschulen aufzustocken, damit Forschung weniger von Drittmitteln abhängig ist und Gutachter entlastet werden (auch durch eine gezielte Nachwuchsrekrutierung); im April nun bezog sie zur Replikation von Forschungsergebnissen Stellung mit dem Ergebnis, dass es „neben individuellem Fehlverhalten für das Qualitätsproblem von Forschung auch strukturelle Gründe gibt. Das mittlerweile in der Wissenschaft erreichte Gewicht von quantitativ parametrisierenden Steuerungs-, Bewertungs- und Gratifikationssystemen wirkt sich auf die Forschung als gestiegener (und weiter steigender) Wettbewerbs- und Beschleunigungsdruck aus.“ Momentan will sich das Selbstverwaltungsorgan der deutschen Wissenschaft in der duz nicht äußern, da eine eigene Wortmeldung zum Peer Review vorbereitet werde, wie es heißt.

Ausdauer fördern

Der Literaturwissenschaftler Prof. Dr. Peter-André Alt, Präsident der Freien Universität Berlin, kritisierte 2014, dass den Geisteswissenschaften das in den Naturwissenschaften übliche „halbwegs objektive“ Double-Blind-Peer-Review fehle, zugleich warnte er vor „neuen Dimensionen eines ‚Forschungsmülls‘“, der daraus resultiere, dass kein „durchgreifendes Prüfsystem“ bei Online-Publikationen entstehe, und lobte, dass renommierte Universitätsverlage in Amerika „nicht unerhebliche Honorare“ für geisteswissenschaftliche Monografien zahlten. Nach Alt „kein schlechter Trend, denn er fördert Tugenden, die im Wissenschaftsbetrieb verloren zu gehen drohen“, darun-

ter „intellektuelle Geduld und Ausdauer, geistige Unabhängigkeit gegenüber Moden“. Gegenüber Open-Peer-Review-Verfahren hat er auch heute noch Vorbehalte: „Zwar sollten nach den Regeln dieses Ansatzes solche Texte, die den Qualitätskriterien nicht standhalten, wieder aus dem Netz verschwinden, aber die Unübersichtlichkeit nimmt zu.“ In Abgrenzung hierzu liegt für ihn eine „klare Differenzierung“ im Double-Blind-Verfahren, das „die Ri-

nur formale Standards durch Peers prüfen zu lassen und die inhaltliche Auseinandersetzung „nach der Veröffentlichung im Fach und damit wissenschaftsöffentlich“ zu führen.

Honorare für Gutachter

Chemieprofessor Dr. Helmut Schwarz, Präsident der Alexander von Humboldt-Stiftung, der beim Neujahrsempfang 2016 mit Blick auf Forscher, die um 1000 Gutachten pro Jahr gebeten würden, vor einem Qualitätsverlust warnt, „der sich zu einem Kollateralschaden auswachsen könnte“, sieht heute Reformpotenzial darin, Gutachter beim Lehrdeputat zu entlasten, aber auch, den durch zeitlich begrenzte Projektfinanzierungen entstehenden Aufwand zu senken durch eine „langfristiger angelegte Finanzierung, die mehr Kredit für und Vertrauen in bereits erbrachte Leistungen honoriert“.

Um die Wertschätzung der Gutachter sorgt sich auch der Wirtschaftswissenschaftler Ulrich Berger, er setzt darauf, Gebühren für Gutachten zu erheben. „Es ist ein kleines Wunder, dass der Peer-Review-Prozess überhaupt so gut funktioniert, wie er es tut. Er könnte aber meiner Einschätzung nach deutlich besser funktionieren, wenn man die oft aufwendige Gutachtertätigkeit über bescheidene 'submission fees' honorieren würde.“

Auf der Suche nach neuen Wegen der Leistungsbeurteilung fällt immer wieder auch der Begriff „Altmetrics“. Während Osterloh in diesen aus den sozialen Medien abgeleiteten Maßzahlen kein Instrument sieht, um eine fairere Qualitätsbeurteilung zu erreichen, verkündet im Mai die EU-Expertenkommission für Altmetrics, sie sähe traditionell zitationsbasierte Metriken (Impact Factor), aber auch nutzungsorientierte Metriken (Altmetrics) als „wertvolle Ergänzung zu Peer Review“ an. Für Kommissionsmitglied Prof. Dr. Isabella Peters liegt ein Vorteil von Altmetrics darin, dass sie „im Gegensatz zu Zitationen, die sich vornehmlich auf wissenschaftliche Publikationen beziehen und nur den Einfluss auf andere Autoren und Autorinnen anzeigen, auch die Nutzung wissenschaftlicher Inhalte durch andere Öffentlichkeiten, wie etwa die Politik, Presse oder interessierte Laien reflektieren“. Wenngleich Metriken Peer Review ergänzen könnten, so spielten bei allen Analysen wissenschaftlicher Leistung Werturteile eine Rolle, die entweder dem, was gezählt werden kann, oder aber disziplinären Konventionen und Schulen einen Vorzug gäben. Peters Fazit: „Der jetzt stattfindende Paradigmenwechsel zu Open Science ist ein optimaler Zeitpunkt, zu rekapitulieren, was schief gelaufen ist.“

Dr. Pascale Anja Dannenberg ist duz-Redakteurin.

„Zeit, zu rekapitulieren, was schief gelaufen ist“

siken der Patronage und Kartellbildung reduziert, Anonymität sichert und Entscheidungen unabhängig von Reputation und Karrierestatus ermöglicht“.

Auf Vorab-Begutachtungen setzt auch der Althistoriker Dr. Christoph Lundgreen, Mitglied im Präsidium der Jungen Akademie. Peer Review sei „alternativlos“; das Verfahren werde jedoch durch „zu viele Anträge und Artikel, ihrerseits Folge der schlechten Grundfinanzierung der Universitäten“, erschwert. Zu überlegen sei, bei Anträgen die „Qualität schon geleisteter Arbeit“ zu berücksichtigen, bei Publikationen

