

Digitale Resonanz Neue Bewertungskulturen fordern die Wissenschaft heraus

Martina Franzen

Summary: Evaluation practices are shaping the digital economy. Users are constantly encouraged to evaluate products and services in order to improve them and, in doing so, themselves are becoming objects of evaluation. These reciprocal evaluation practices are the main characteristics of the Web 2.0 culture, which find their analogue precursor in the science system and the peer review system. The article explores the implications of digitalization based on the case of Altmetrics, i.e. tools that measure the web impact of scientific outputs.

Kurz gefasst: Bewertungspraktiken aller Art prägen die digitale Ökonomie. Jeder Kunde ist angehalten, Produkte und Services zu bewerten, um sie zu optimieren, aber auch er selbst wird zum Bewertungsobjekt. Das Prinzip eines reziproken Bewertungssystems als Kennzeichen der Web 2.0-Kultur findet seinen analogen Vorläufer im Wissenschaftssystem und im Peer Review. Am Beispiel von Altmetrics, also Tools zur Erfassung von Netzreaktionen auf wissenschaftliche Publikationen, werden die Folgen des Wandels durch die Digitalisierung näher bestimmt.

„Wie zufrieden sind Sie mit unserem Service? Hat das Produkt Ihre Erwartungen erfüllt?“ Nahezu jede Transaktion im digitalen Raum ist heutzutage gekoppelt an die Aufforderung zum bewertenden Feedback. Die Sicht des Nutzers beziehungsweise Kunden wird eingeholt, um Produkte zu optimieren oder aber das Marketing zu verbessern. Neben Ratings auf unterschiedlichen Skalen sind häufig inhaltliche Erläuterungen in Form von Rezensionen gefragt, die wiederum selbst Gegenstand der Bewertung werden können, um die multiplen Bewertungen zu hierarchisieren: „War diese Rezension hilfreich?“ Bei positiver Antwort steigt nicht zuletzt der kalkulierte Reputationswert des Rezensenten, der bis in die Riege der Top-Rezensenten aufsteigen kann, was als Anreizsystem für weitere Bewertungen fungiert.

In einer Episode der Netflix-Serie „Black Mirror“ wird die allgegenwärtige Jedermann-Bewertung als neue Form sozialer Kontrolle detailreich entfaltet. Gezeichnet wird das düstere Bild einer auf einem gegenseitigen Bewertungssystem beruhenden Gesellschaft, in der das Individuum in jeder sozialen Interaktion nur noch darum buhlt, möglichst viele *Likes* einzufangen, um den eigenen Reputationscore zu erhöhen. Bewertet wird jede Begegnung oder Serviceleistung. Die Bewertung findet von beiden Seiten statt, in Echtzeit und online für alle einsehbar. Die Protagonistin Lacie führt uns anschaulich vor, wie der personenzentrierte Score das individuelle Verhalten kontrolliert, sobald das Erreichen eines hohen numerischen Werts nicht nur einem narzisstischen Bedürfnis folgt, sondern den sozialen Status bestimmt. Ein scheinbar kleiner Unterschied zwischen 4.2 und 4.5 auf der Punkteskala wird für die ehrgeizige Lacie zur unüberwindbaren Hürde auf ihrem Weg zum sozialen Aufstieg.

Ähnlich wie der Figur Lacie aus „Black Mirror“ ein Therapeut zur Seite steht, der sie strategisch darin berät, wie sich ihr Reputationscore in kürzester Zeit erhöhen lässt (um eine der begehrten Wohnungen zu bekommen), existieren auch für den Wissenschaftsbereich zahlreiche teils ernsthafte, teils zynische Ratgeber zur Maximierung der eigenen Resonanzquote. Ein Beitrag auf der Network Website Academia.edu unter dem vielversprechenden Titel „How to Increase Your Papers' Citation and H Index“ erreicht aktuell rund 50.000 Views. Die ironische Strategieempfehlung lautet: Erhöhe die Selbstzitationen massiv, um die nötige Aufmerksamkeit für dich zu erreichen. Dieser kurze Text, der die ethisch fragwürdigen Publikationspraktiken eines indonesischen Physikprofessors und mutmaßlichen Fachkollegen des Autors offenlegt beziehungsweise anzeigt, illustriert ein in der Wissenschaft viel diskutiertes Problem indikatorenbasierter Leistungsmessung: Die Orientierung an Nutzungsstatistiken in der Evaluationspraxis setzt *Gaming*-Aktivitäten in Gang, die das meritokratische Prinzip der Leistungsgerechtigkeit nach wissenschaftlicher Qualität unterhöheln und zu Vertrauensverlusten führen können.

Über die Frage, inwiefern sich qualitative Eigenschaften sinnvoll in quantitative Maßzahlen überführen lassen, wird speziell in der Wissenschaft seit Längerem gestritten. Neben dem klassischen Instrument der qualitativen Bewertung, dem *Peer Review System*, haben mit der Einführung des *New Public Management* im Hochschulbereich quantitative Indikatoren Einzug gehalten, etwa wenn es um die Zuteilung von Fördergeldern geht.

Aus der Evaluationsforschung ist seit Längerem bekannt, dass jede Output-Kontrolle über Kennziffern mit einer Vernachlässigung der Inhalte einhergeht. Eindimensionale Indikatoren können Zielkonflikte im System hervorrufen. Mar-

shall W. Meyer und Vipin Gupta haben es als „Performanz-Paradox“ beschrieben, wenn Indikatoren nicht mehr dazu taugen, gute und schlechte Performanz zu unterscheiden. Im Bereich zitationsbasierter Indikatoren in der Wissenschaft wie dem h-Index oder dem Journal Impact Factor vollzieht sich *Gaming the System* auf verschiedenen Ebenen: Es betrifft Autoren, Herausgeber wie Verlage. Sanktionsmaßnahmen gegenüber ethisch fragwürdigen Publikations- oder Zitationspraktiken sind bekanntermaßen beschränkt.

Die Digitalisierung fügt der Orientierung an Resonanzquoten in der Wissenschaft eine neue Dimension hinzu: Mit dem Kunstwort Altmetrics werden Methoden bezeichnet, die ein möglichst weites Spektrum an Netz-Reaktionen auf Veröffentlichungen erfassen sollen. Dem Konzept liegt der Anspruch zugrunde, die Wissenschaft demokratisieren zu wollen, indem ein offenes und gerechteres System der Leistungsbewertung geschaffen wird. Dies war jedenfalls der Tenor des Altmetrics-Manifesto, das im Jahr 2010 den diskursiven Grundstein für die weitere soziotechnische Entwicklung legte.

Altmetrics bilden sowohl das ganze Spektrum von *research outputs* wie Zeitschriftenartikel, Bücher, Datensets, Blogposts, Foliensätze als auch deren multiple Verwendungsweisen unterhalb der Zitationsschwelle (unter anderem Bookmarks, Downloads, Views) ab. Statt eines journal- oder eines personenzentrierten Indikators haben wir es bei Altmetrics mit einem ereigniszentrierten Bewertungstool zu tun. Die webbasierten Messungen beziehen sich auf ein beliebig erweiterbares Repertoire an Quellen, statt sich wie die herkömmlichen Indikatoren auf die Zitationsstatistiken von Quellzeitschriften zu beschränken, wie sie im Web of Science oder in Scopus gelistet sind. Nimmt man den Dienstleister Altmetric.com als Beispiel, so zählen unter anderem soziale Netzwerke wie Facebook, Microblogging-Dienste wie Twitter, Videoplattformen wie Youtube genauso wie internationale und nationale Medientitel zu den definierten Quellen für die automatisierte Erfassung von Impact. Altmetric.com, ein Geschäftsbereich von Digital Science als Zweig von Macmillan Publishers Ltd., ist jedoch vor allem für seinen Artikel-Score bekannt. Der auf einem nicht offengelegten Algorithmus basierende Altmetric-Score wird in Form von sogenannten Badges ausgewiesen. Besonders beliebt ist der Altmetric Donut – ein Ring, dessen Färbung über die Art des erzielten Impacts, also über Quellen (blau für Twitter, rot für Presse etc.) informiert. Eine gelungene technische Spielerei, könnte man meinen, aber für die Wissenschaft bedeutungslos. Die Sicht der Proponenten, deren Kreis sich aus Teilen der Wissenschaft, aus IT-Branche, Bibliothekswesen und Wissenschaftsverlagen zusammensetzt, ist jedoch eine andere: Altmetrics sollen zum Katalysator des Umbaus des akademischen Reputationssystems werden. Wie aber sollten Tweets oder Facebook-Likes Auskunft über wissenschaftliche Qualität oder Relevanz geben können?

Während die methodologisch entscheidende Frage unbeantwortet bleibt, schreitet die flächendeckende Implementation von Altmetric-Tools in die digitalen Publikationsinfrastrukturen weiter voran. So haben die großen internationalen Verlage Elsevier, Wiley oder Springer sowie die Spitzenzeitschriften Science und Nature sie bereits in ihr Portfolio integriert. Auch das soziale Netzwerk ResearchGate setzt auf Altmetrics auf Basis der erfassten Publikationsdaten und der personalisierten Nutzungsstatistiken – mit dem einen Unterschied, dass es zusätzlich einen personenzentrierten Score ausweist. Dieser Score setzt Wissenschaftler in Beziehung zueinander. Transparenz wird dadurch hergestellt, dass dem ResearchGate-Mitglied und all seinen Rezipienten genau vorgerechnet wird, in welchem Perzentil der individuelle Score angesiedelt ist. Der Score ist ein kumulativer Wert, der jedoch leichten und zum Teil irritierenden Schwankungen selbst nach unten ausgesetzt ist. Wöchentlich werden die Nutzer über die statistische Kurve der Nutzung eigener oder anderer Beiträge informiert. Als Motivationsstruktur wurde darüber hinaus ein klassisches Element aus der Spielebranche in die wissenschaftliche Kommunikation transferiert, und zwar die Markierung des Eintritts in ein neues Level, dessen Basis hier die Erreichung eines Schwellenwerts an Zitationen oder Klicks oder ein Spitzenplatz im Institutsranking darstellt. Wie in digitalen Spielumgebungen üblich, werden auch bei ResearchGate Hilfestellungen angeboten, um den individuellen Score zu erhöhen, zum Beispiel: „Boost your stats by adding more research.“



Martina Franzen ist wissenschaftliche Mitarbeiterin der Forschungsgruppe Wissenschaftspolitik. Die promovierte Soziologin forscht über das Verhältnis von Wissenschaft, Medien und Öffentlichkeit, aktuell vor allem zu den Folgen der digitalen Wende der Wissenschaft. (Foto: David Ausserhofer)

martina.franzen@wzb.eu

Über ein solches Anreizsystem werden die digitalen Plattformen ResearchGate und Academia.edu mit Daten jeder Art sukzessive angereichert. Big Data bildet somit die Basis ihres Geschäftsmodells, dessen Konturen sich bislang nur verschwommen abzeichnen. Im letzten Jahr führte Academia, das amerikanische Pendant des deutschen Start-ups ResearchGate, einen kostenpflichtigen Premium-Account mit einer jährlichen Mitgliedsgebühr von 99 US-Dollar ein. Noch ist das Premium-Feature inhaltlich kaum unterscheidbar vom Freemium Account – der einzige Unterschied betrifft den Zugang für Autorinnen, mit dem sie einen detaillierten Überblick über die Art der Nutzung und die Merkmale des Nutzers beziehungsweise Lesers einzelner Beiträge bekommen, aufgeschlüsselt nach Personen und aggregiert nach Institutionen, Ländern etc.

War auf den digitalen Plattformen bislang der Wissenschaftler als Autor Gegenstand der Bewertung, wird nun zusätzlich der Wissenschaftler als Leser in Bezug auf sein individuelles Nutzungsverhalten vermessen. Ob sich eine Bezahl-schranke in der wissenschaftlichen Kommunikation durchsetzen lässt, ist vor dem Hintergrund der Open-Science-Bewegung mehr als fraglich. Doch trifft die Offerte, mehr über den Leserkreis seiner Publikationen zu erfahren, auf ein genuines Bedürfnis des Wissenschaftlers, der sich nicht wie die Literatin über den Abverkauf von Büchern ein Bild über die Reichweite seiner Veröffentlichungen machen kann. War über herkömmliche Zitationsmaße bislang quasi nur die Spitze des Eisbergs erkennbar, machen Altmetrics nun die ganze Breite an Rezeptionsweisen diesseits der formalen Zitierung in akademischen Fachzeitschriften einsehbar und zum Gegenstand der Bewertung von Wissenschaft. Dies befriedigt ein narzisstisches Bedürfnis von Wissenschaftlern und bietet möglicherweise einen informationellen Mehrwert für die institutionelle Forschungsevaluation. Die entscheidende Frage, was Altmetrics eigentlich genau messen, ist allerdings bislang unbeantwortet.

Der dominante Forschungsansatz der Bibliometrie, vergleichende empirische Studien zwischen Altmetriken jeder Art und Zitationsraten anzustellen, hilft in der Sache kaum weiter. Zwar lässt sich theoretisch die Zitation als Form sozialer Anerkennung wissenschaftlicher Leistungen fassen. Wenn die Evaluationspraxis aber von der schieren Anzahl der Zitationen auf Qualitätsunterschiede wissenschaftlicher Beiträge schließen will, bedeutet das einen Kurzschluss zwischen Resonanz, also Popularität, und Qualität. Auf der Basis meiner eigenen Arbeit zur Medialisierung von Wissenschaft formuliere ich eine andere Annahme: Altmetrics – wie auch Zitationsraten – zeigen in erster Linie Popularität an. Hohe Resonanzquoten können mit wissenschaftlicher Qualität zusammenfallen, ebenso beeinflusst aber sind sie durch Nachrichtenfaktoren wie Unterhaltung, Skandale oder Prominenz. Die Aussagefähigkeit von Altmetrics (und Zitationsraten) lässt sich demnach primär auf die Messung von Marketingenerfolg reduzieren. Marketingenerfolg im Sinne eines hohen erzielten Impacts kann auf die besondere inhaltliche Qualität des wissenschaftlichen Beitrags verweisen, doch sind die politische Brisanz oder Aktualität des Themas, die Prominenz des Autors oder auch einfach die gut platzierte Werbung genauso impactförderlich.

Ein anschauliches Beispiel für die These, dass wissenschaftliche Werte mit Nachrichtenwerten im Bereich der Impact-Messung in Konflikt geraten können, liefert das Jahresranking der Top-100-Artikel per Altmetric Score: Den Spitzenplatz für 2016 nimmt ein Artikel aus dem renommierten *Journal of the American Medical Association* ein. Der Autor ist nicht etwa ein medizinischer Forscher, sondern der damals amtierende Präsident der Vereinigten Staaten von Amerika, Barack Obama höchstpersönlich, der zum Thema der US-Gesundheitsreform schrieb. Dass eine solche Honorierung des höchsten Altmetric Scores wenig mit wissenschaftlichen Relevanzkriterien zu tun hat, ist offensichtlich. Umso verwunderlicher ist, dass Altmetrics bislang in der Wissenschaftsgemeinschaft kaum kritisch hinterfragt werden.

Paul Wouters und Rodrigo Costas haben das Altmetrics-Konzept als „Narzissmustechnologie“ bezeichnet. Hieraus erklärt sich vermutlich auch sein rasanter Aufstieg. Die Frage ist, ob es seinen hohen Zuspruch auch dann behält, wenn es tatsächlich zu einer „Kontrolltechnologie“ mutieren sollte. Altmetrics sind zwar noch nicht offiziell als Evaluationsinstrument in die institutionelle Leistungsbe-

wertung eingeführt, doch scheint ihre Implementation zum Beispiel für die Messung gesellschaftlichen Impacts eher eine Frage der Zeit. Jede Art der Leistungsmessung aber löst reaktives Verhalten aus und bleibt für das System nicht folgenlos. Was als *Gaming the System* kritisiert wird, kann auch als erfolgreiche Anpassung an fehlgeleitete Indikatoren gelesen werden. Das Spiel ist alt – beim Impact-Faktor betraf Gaming vor allem die Zeitschriften (über redaktionelle Auswahl und PR), bei Altmetrics kommen nun die Autoren selbst ins Spiel. Für sie geht es darum, erfolgreiches Reputationsmanagement zu betreiben und über Eigenwerbung in sozialen Medien, aktive Bewertungen der anderen oder passgenaue Zielgruppenkommunikation dafür zu sorgen, die eigenen Klickzahlen stetig zu erhöhen. Das Streben nach maximaler Reichweite erfordert jedoch andere Mittel und ist nicht genuines Ziel der wissenschaftlichen Arbeit. Es ist vielmehr eine Antwort auf die Bedingungen der Aufmerksamkeitsökonomie im digitalen Zeitalter – mit allen möglichen Folgen für die Qualität des produzierten Wissens im Horizont der Erkenntnisbildung.

Literatur

Black Mirror, Staffel 3, Episode 1 („Nosedive“), erstmalig ausgestrahlt bei Netflix im Oktober 2016.

Espeland, Wendy N./Sauder, Michael: „Rankings and Reactivity: How Public Measures Recreate Social Worlds“. In: *American Journal of Sociology*, 2007, Vol. 113, July, pp. 1–40.

Franzen, Martina: *Breaking News. Wissenschaftliche Zeitschriften im Kampf um Aufmerksamkeit*. Baden–Baden: Nomos 2011.

Franzen, Martina: „Der Impact Faktor war gestern. Altmetrics und die Zukunft der Wissenschaft“. Themenheft: *Der impact des impact factors*. In: *Soziale Welt*, 2015, Jg. 66, H. 2, S. 225–242.

Meyer, Marshall W./Gupta, Vipin: „The Performance Paradox“. In: *Research in Organizational Behavior*, 1994, Vol. 16, pp. 309–369.

Osterloh, Margrit: „Governance by Numbers: Does It Really Work in Research?“ In: *Analyse und Kritik*, 2010, Bd. 32, H. 2, S. 267–283.

Wouters, Paul/Costas, Rodrigo: „Users, Narcissism and Control – Tracking the Impact of Scholarly Publications in the 21st Century“. Utrecht: SURFfoundation 2012. Online: <http://research-acumen.eu/wp-content/uploads/Users-narcissism-and-control.pdf> (Stand 15.01.2017).