

Forum

BODO RÖDEL

Open Science und unser Bild von Wissenschaft

Unter Open Science wird das Bestreben verstanden, sowohl die Ergebnisse wissenschaftlicher Forschung als auch den Forschungsprozess selbst offen zugänglich zu gestalten. Die Idee des Open Science wird häufig durch sechs grundlegende Prinzipien charakterisiert: Die Methodologie wissenschaftlicher Forschung ist dokumentiert; die verwendete Technologie liegt sowohl mit Blick auf die Soft- als auch auf die Hardware offen; alle verwendeten Daten sind frei zugänglich; alle Ergebnisse sind so publiziert, dass sie offen zugänglich sind; der Prozess der Qualitätssicherung ist transparent und schließlich: Lehr-Lern-Materialien sind frei und ohne Einschränkungen verfügbar. Der Begriff des Open Science schließt also die Begriffe Open Access und Open Educational Resources mit ein.

Der Begriff „Open“ bedarf noch einmal einer gesonderten Betrachtung. Es gibt einige dezidierte Versuche, den Begriff „Open“ zu definieren. Der elaborierteste findet sich wohl unter opendefinition.org: „Knowledge is open if anyone is free to access, use, modify, and share it — subject, at most, to measures that preserve provenance and openness.“ (Stand 14.2.2015). Im Detail bedeutet das u. a.: Der Zugang zu wissenschaftlicher Forschung soll offen (nicht zu verwechseln mit frei im Sinne von kostenlos) sein. Auch die Weiterverbreitung ist so zu gestalten, dass sie offen erfolgen kann – dies gilt auch für Modifikationen der Ursprungsquelle. Allerdings kann die Namensnennung zu einer Bedingung der Weitergabe gemacht werden. Technische Einschränkungen sind zu vermeiden. Nachnutzungen sind ggfls. durch andere Namen oder Versionsnummern kenntlich zu machen.

Die Idee des Open Science bliebe allerdings oberflächlich erklärt, würde sie nur auf die offene Weitergabe von Daten und Ergebnissen reduziert (VON CASTELL 2015). Es stellt sich vielmehr auch die Frage, inwieweit sie das verändert, was als gute wissenschaftliche Praxis, mithin Wissenschaft als systematische Methode, beschrieben wird (siehe zum Beispiel DFG 2013).

Für viele Forschungszweige kann die Informationsverarbeitungstechnik (IT) als das zentrale Werkzeug der Forschung gelten, da sie von der Verarbeitung großer Datenmengen leben. Die moderne IT hat aber nicht nur die Wissenschaft, sondern auch die Gesellschaft verändert – Stichworte sollen hier genügen: Google, Facebook, NSA. Die zunehmende Nutzung und die Abhängigkeit von digitalen Daten („Digitalisierung“) ist also auf zweifache Weise „normal“: Zum einen für die Gesellschaft direkt, zum anderen als Bestandteil der Wissenschaft. So stellt PETER HINSEN fest: „We’re living through an age where technology has fundamentally changed society. (...) We’re living through an age where technology has become a commodity, where digital has become ‚normal‘.“ (HINSEN 2015).

Diese Digitalisierung führt – so HINSEN – zu einer Verkürzung der Zeitspanne, in der Information relevant bleibt. Sind dies bei Büchern vielleicht Jahre, haben Tweets nur eine Relevanz von wenigen Minuten. Die Aufnahme und Verarbeitung von Informationen beschleunigt sich also immer mehr. Große Organisationen, wie beispielsweise auch die

Wissenschaftsorganisationen, müssen auf diese Entwicklung reagieren. Die Reaktionszeit ist dabei i. d. R. langsamer als das, was von außen an Herausforderungen an diese Organisationen herangetragen wird – oder quillt Ihr E-Mail-Postfach nicht über? Die Digitalisierung ist für einen Prozess verantwortlich, in dem durch die zunehmende Beschleunigung der Wissensproduktion die meisten von uns Getriebene sind. Es geht nicht mehr um die Frage, ob wir solch eine Digitalisierung gut oder schlecht finden, sie ist bereits „the new normal“, so HINSSEN. „But we're probably only halfway in this transition, halfway into the New Normal where we will have to see which business models will survive, and which companies will remain relevant in the New Normal. (...) Information now flows faster than ever before. Customers have access to more information than ever before, and information becomes 'stale' faster than ever before.“ (HINSSEN 2015).

Die Verarbeitung von Informationen findet nicht irgendwann in der Vergangenheit statt, sondern jetzt, im Augenblick. HINSSEN nennt dies „the era of now“ (HINSSEN 2010). Diese Ära ist durch zwei Effekte gekennzeichnet: Die Verarbeitung und der Zugang zu Informationen muss schnell und jederzeit möglich sein. Ansonsten muss der Zugang nicht perfekt sein – aber schnell. Dauert das Laden einer Website zu lange, klicken wir weg. Die IT reagiert auf diese zwei Ansprüche zum einen mit der Entwicklung von Cloud Computing – der Zugriff auf Daten in der Cloud ist prinzipiell von überall aus möglich. Die zweite Bewegung ist Big Data. Dieses Schlagwort bezeichnet seit 2012 das Sammeln und Auswerten großer Datenmengen. Diese Datenmengen sind zu groß und/oder verändern sich zu schnell, als dass sie noch an einem PC von einer Person verarbeitet werden könnten (DUMBILL 2015).

Ist das aus der privaten Perspektive eher ein Problem von Prioritäten, so beschäftigt sich Wissenschaft zunehmend mit der Auswertung großer Datenbestände: Text- und Data-Mining werden die Gesellschaft und damit Wissenschaft in Zukunft entscheidend verändern: „Businesses use data and text mining to analyse customer and competitor data to improve competitiveness; the pharmaceutical industry mines patents and research articles to improve drug discovery; within academic research, mining and analytics of large datasets are delivering efficiencies and new knowledge in areas as diverse as biological science, particle physics and media and communications.“ (McDONALD/KELLY 2015). Eine der zentralen Voraussetzungen für Text- und Data-Mining ist eine open access-Veröffentlichung der Forschungsergebnisse.

Vielleicht ist anders als durch automatisierte Verfahren auch keine Bewältigung der Flut an Daten und Veröffentlichungen mehr möglich. Das Feld der Berufsbildungsforschung scheint hier noch eher am Anfang zu stehen. Aber schon jetzt verzeichnet die vom Bundesinstitut für Berufsbildung (BIBB) zur Verfügung gestellte Literaturdatenbank Berufliche Bildung (LDBB) alleine für das Jahr 2014 rund 1.400 neue Einträge. Der Datenreport zum Berufsbildungsbericht speist sich aus über 40 einzelnen Datenquellen. Das Forschungsdatenzentrum des BIBB hält rund 246 Datensätze mit im Schnitt 122 Variablen vor. Diese Ansammlung von Daten und Publikationen genügt natürlich bei weitem noch nicht den Maßstäben von Big Data. Selbst in dem engen Feld der Berufsbildungsforschung dürfte es für eine einzelne Person aber schwierig werden, den Überblick zu behalten.

Was aber bedeutet es für die Wissenschaft, wenn der Forschungsprozess auf IT basiert und damit der Prozess selbst immer mehr an Geschwindigkeit gewinnt, gleichzeitig Wissenschaftler aber mit Blick auf ihre Karriere gezwungen sind, zu publizieren? Schon heute ist zu vermuten, dass nicht nur die Qualität, sondern auch die Quantität an Veröffentlichungen die wissenschaftliche Karriere eines Forschers entscheidend mitbestimmt und diese sogar vorhersagbar macht (VENCE 2015). Beides scheint dabei in einem Spannungsverhältnis zu stehen, in der die Qualität wohl öfters als vermutet zu kurz zu kommen scheint (JOHNSON 2015). Sollten wissenschaftliche Journals durch das peer-review-Verfahren eigentlich Garanten für Qualität sein, so ist zu befürchten, dass sie eher interessante Ergebnisse - die aber nicht notwendigerweise auch die richtigen Ergebnisse sein müssen - den Vorzug geben. Daher wird ihre Rolle durchaus kritisch gesehen (SCHEKMAN 2015). Auch ist die Wahrscheinlich-

keit, dass ein Artikel publiziert wird, höher, wenn er positive Ergebnisse beinhaltet. Dies hat Folgen: „A system that disfavors negative results not only distorts the scientific literature directly, but might also discourage high-risk projects and pressure scientists to fabricate and falsify their data.“ (FANELLI 2012, S. 891).

Die Qualität von Forschungsergebnissen und die Überprüfung dieser Ergebnisse hängen von den verwendeten Daten ab und von der Möglichkeit, Zugriff auf diese Daten zu bekommen. Open Science und der damit verbundene Anspruch, nicht nur die Forschungsergebnisse, sondern auch alle dahinterliegenden Daten transparent zugänglich zu machen, wird damit vor dem Hintergrund der zuvor geschilderten publish-or-perish-Problematik zu einer Frage der Kommunikation, aber auch und wesentlich zu einer Frage der Qualitätssicherung wissenschaftlicher Forschung.

Konsequent weitergedacht ergeben sich durch Open Science schließlich alternative wissenschaftliche Methoden. „Not only will our knowledge increase but the way science is done will change in profound ways. A new ecology of science will crystallize, one that shifts from the insular and closed structures of academic, industrial, and military research toward open models based on social connection, data commons, and democratized tools and technology. We will create unprecedented opportunities for collaboration and resource sharing between large organizations, communities, and individuals. (...)“ (PESCOVITZ 2011) Es bleibt damit abzuwarten ob die Digitalisierung nicht nur die Gesellschaft und die Publikationsformen von Wissenschaft, sondern auch den Forschungsprozess an sich verändern wird.

Literatur

- DEUTSCHE FORSCHUNGSGEMEINSCHAFT: Vorschläge zur Sicherung guter wissenschaftlicher Praxis. Weinheim, 2013
- DUMBILL, E.: What is big data?. <http://radar.oreilly.com/2012/01/what-is-big-data.html> (Stand: 17.2.2015)
- FANELLI, D.: Negative results are disappearing from most disciplines and countries. *Scientometrics*, 2012 90:891–904; DOI 10.1007/s11192-011-0494-7
- HINSEN, P.: The New Normal. Tielt/Belgien, 2010
- HINSEN, P.: Topics. http://www.peterhinssen.com/keynotes_and_workshops/topics (Stand: 17.2.2015)
- JOHNSON, G.: New Truths That Only One Can See. http://www.nytimes.com/2014/01/21/science/new-truths-that-only-one-can-see.html?_r=2 (Stand: 17.2.2015)
- MCDONALD, D.; KELLY, U.: The Value and Benefits of Text Mining to IK Further and Higher Education. <http://www.jisc.ac.uk/sites/default/files/value-text-mining.pdf> (Stand: 17.2.2015)
- PESCOVITZ, D.; GORBIS, M.; WALDMAN, A.: A multiverse of Exploration – The Future of Science 2021. Palo Alto, 2011
- SCHKEMAN, R.: How journals like Nature, Cell and Science are damaging science. <http://www.theguardian.com/commentisfree/2013/dec/09/how-journals-nature-science-cell-damage-science> (Stand: 17.2.2015)
- VENCE, T.: Can Publication Records Predict Future PIs?. <http://www.the-scientist.com/?articles.view/articleNo/40118/title/Can-Publication-Records-Predict-Future-PIs/> (Stand 17.2.2015)
- ZU CASTELL, W.: Open Science – Mehr als nur eine Frage der Kommunikation. https://www.youtube.com/watch?v=6NrdD_RqZs (Stand 17.2.2015)

Dr. Bodo Rödel, Bundesinstitut für Berufsbildung, Leiter des Arbeitsbereichs „Publikationsmanagement/Bibliothek“, Robert-Schuman-Platz 3, 53175 Bonn, E-Mail: roedel@bibb.de