

Alle Rechte beim Urheber.  
Abdruck nur gegen Belegexemplar.

### **Offene Fragen zu Open Data**

Sie ist so unausweichlich wie ungenügend, die Offenlegung von Forschungsdaten. Unausweichlich, weil aus der anschwellenden Datenflut mit immer höherer Treffsicherheit geschlossen werden kann auf die Daten, die zurückgehalten werden. Und ungenügend, weil Daten allein nichtssagend sind - entscheidend ist bei Daten immer der Kontext, wie und wozu sie gewonnen wurden; und - im Wissenschaftsbetrieb - wer sie gewonnen hat. Sonst geht der Nobelpreis ja möglicherweise an den Falschen, wer auch immer das sein könnte!

Aufgrund dieser - nicht auf Wissenschaftler beschränkten - selbstverliebten Eitelkeit heraus, die in der Renaissance als Amour propre gegeißelt wurde, war und ist die Bereitschaft von Wissenschaftlern, ihre Daten offenzulegen, nur sehr gering ausgeprägt. Sie steigt im selben Maße, wie das Offenlegen der eigenen wissenschaftlichen Laufbahn dienlich ist. Dabei folgt die Offenlegung konzentrischen Kreisen, die sich weiten vom Kollegen über den Förderer bis hin zur allgemeinen Öffentlichkeit.

Das differenzielle Teilen von Forschungsdaten wird zudem beeinflusst von der gesellschaftlichen Erwartungshaltung. Je schlechter diejenigen Forscher angesehen werden, die ihre Daten zurückhalten, umso größer wird der Erwartungsdruck, die Daten herauszurücken. So zwar, dass man mit den Daten etwas anfangen kann. Beispielsweise einen Versuch wiederholen. Mit anderen Worten: die Daten sollten mit Metadaten versehen sein. Die Veröffentlichung und Nutzung metadatierter Daten würde vereinfacht, gäbe es hierfür Standards.

Ausgangspunkt für eine Standardisierung könnte die Bermuda-Vereinbarung aus dem Jahre 1996 sein, in der sich die an der Erforschung des menschlichen Genoms Beteiligten verpflichtet

hatten, die entschlüsselten DNS-Sequenzen in einer öffentlich zugänglichen Datenbank zu hinterlegen. Die Vereinbarung hat die Sequenzierung des Erbguts des Menschen enorm beschleunigt, weshalb die Anhänger von Open Data gerne auf die Katalyse des wissenschaftlichen Fortschritts hinweisen.

Der Hinweis ist insofern erstaunlich, als man sich die Augen reiben möchte, dass er überhaupt nötig scheint. Wissenschaft ist ohne Öffentlichkeit undenkbar. Ein Experiment ist in dem wesentlichen Sinne öffentlich, dass es von jedem und überall wiederholt werden kann, so dass seinen Ausgang jeder mit eigenen Augen bezeugen kann. Wissenschaftliches Wissen ist öffentliches Wissen und gerade nicht das einer esoterischen Kaste. Jedes Hindernis, das sich, wie die Geheimnistuerei um Daten, der Wiederholung eines Versuchs entgegenstellt, ist kontraproduktiv und dazu angetan, das Vertrauen in Wissenschaft und Forschung zu unterminieren.

Ein solcher Vertrauensverlust droht allerdings auch von der Gegenseite: vom Daten-Missbrauch. Im Dataversum der Digitalen Gesellschaft sind kaum noch Daten vorstellbar, die nicht in eine - ausbeutbare - Beziehung gesetzt werden können zu einzelnen Personen, mögen sie von noch so fernen oder abstrakten Objekten gewonnen sein. Die Wissenschaft ist daher gut beraten, wenn sie im Dialog mit Öffentlichkeit und Politik mögliche Missbräuche qualifiziert und Vorkehrungen gegen diese trifft. Insoweit ist Open Data auf eine fortwährende Risikofolgenabschätzung algorithmisierter Datenauswertung angewiesen.

Um sich den problematischen Knoten im Datennetz zuzuwenden zu können, muss das Netz in den Fokus gerückt werden und nicht ihre Erzeuger. Ohne moderne Staaten, deren Bevölkerung sich in einem Gesellschaftsvertrag darauf verständigt hat, wissenschaftlich Einrichtungen zu unterhalten, würden die Urheber wissenschaftlicher Daten geräuschlos durch die Maschen fallen. So als hätte nicht die Wissenschaft höchstselbst dem Urheber die Messe gelesen: Daten zeugen den Urheber, statt dass der Urheber Daten erzeugt. Sozialpsychologisch gewendet: Wissenschaft ist ein Gemeinschaftsprojekt, das nicht nur von Wissenschaftlern getragen wird.

Der Urheber ist mit dem Tod des Autors vor fünfzig Jahren gestorben. Statt also mit einem aufgeblasenen Urheberrecht die Allokation von Privilegien in der Wissenschaft zu betreiben, sollten die Grundrechte vehementer eingefordert und gegen mögliche digitale Übergriffe abgesichert werden. Die Metonymie von den Daten zu ihrem Urheber hält bloß den Mythos am Leben vom Schöpfergott, der statt eines weißen Bartes einen weißen Kittel trägt. Dringlicher wären Antworten auf die Frage, wie wir mit der exponentiell anschwellenden Datenflut umgehen wollen, egal wer sie erzeugt hat.

Dass ausgerechnet von Open Data hierzu Antworten zu erwarten sein könnten, klingt erst einmal paradox. Doch sind mir Antworten auf wissenschaftlicher Basis lieber, als wenn letztlich andere Kräfte die Deutungshoheit gewönner. Wir brauchen gute, reproduzierbare Daten, um eine verlässliche Orientierung für unser Zusammenleben zu finden. Deshalb geht die Deutsche Forschungsgemeinschaft einen Schritt in die richtige Richtung, wenn sie Angaben zur Verwendung der Daten aus einem Forschungsprojekt zu einem Förderkriterium macht. Sie sollte nun auch so konsequent sein und den von ihr Geförderten einen festen Anteil ihrer Fördermittel für eine - brauchbare - Bereitstellung der Daten vorschreiben.