

# Viele Daten, zu wenige Fragen Die Wissenschaft sollte sich dem Kern der Big-Data-Probleme zuwenden

Ranty Islam

In einer mittelgroßen südeinglischen Stadt begannen 2002 mehrere hundert zusammengeschnittene PCs – eine Art Supercomputer für Arme – an einer der grundlegenden Fragen der Menschheit zu arbeiten: Wie ist unsere Welt entstanden? Der Rechner-Cluster an der Universität Oxford brauchte knapp eine Woche für seine Antwort. Sie bestand aus einem mehrere tausend Gigabyte großen Datenpaket.

Illustrieren lässt sich dieser Datenhaufen als ein bizarres Muster aus Knoten, Fäden und Leerräumen. Es zeigt die großen Strukturen unseres Universums, die Vorläufer von Galaxien und Galaxienhaufen, wie diese wenige Milliarden Jahre nach dem Urknall – aber immer noch weit vor unserer Zeit – ausgesehen haben könnten (siehe die Abbildung auf der Rückseite). Der Cluster der Astronomen führte dafür einen Computercode aus, der zwar kompliziert anmutet, aber im Kern nur einen Algorithmus abarbeitet: Er berechnet in unzähligen Iterationen, wie Milliarden von simulierten Materie-Teilchen unter ihrer eigenen Schwerkraft über kleine Zeitabstände aneinander vorbeiziehen oder miteinander verklumpen.

An diesem Experiment konnte ich in meinem früheren Leben als theoretischer Physiker mitarbeiten. So bauten wir im Jahr 2002 das Universum mit Big Data nach – ohne es so zu nennen. Die Physik musste sich dafür Hybris vorhalten lassen. Heute modellieren wir mit Big Data unsere Gesellschaft. Das ist Hybris. Und blauäugig. Zumindest solange es keine kritische Kompetenz gibt, die diese Technologie versteht. Diese schickt sich an, zu einer neuen Kulturtechnik des Menschen zu werden.

Wir kennen die allseits referierten Vorteile, die die digitale Erfassung aller menschlichen Lebensbereiche für die Zukunft von Wirtschaft, Forschung, Medizin, Verbrechensbekämpfung etc. haben soll – etwa über soziale Netzwerke, E-Commerce, digitale Verwaltung oder Kommunikations-Tracking. Doch die kritische Debatte über die gesamtgesellschaftlichen Auswirkungen des Phänomens bleibt unterbelichtet. Klar, wir diskutieren über Pri-

vatsphäre, Überwachung, Netzneutralität, Governance und so weiter. Andere größere, Fragen haben wir noch gar nicht begonnen zu stellen – geschweige denn zu beantworten: Was bedeutet es zum Beispiel, eine Identität zu haben, wenn wir unser Gedächtnis zunehmend durch die Google-Suche ersetzen? Wenn wir unsere Erinnerungen auf Instagram oder unsere Beziehungen mit anderen Menschen auf Facebook oder LinkedIn auslagern? Wenn die Amazon-Kaufhistorie zur Signatur der eigenen Persönlichkeit wird? Am Ende steht die grundlegende Frage, was es unter diesen Umständen bedeutet, ein Mensch zu sein. Oder anders: Wir fragen gerade vor allem, was Menschen mit Big Data machen. Was wir aber wirklich brauchen, ist eine Diskussion darüber, was Big Data mit Menschen macht.

Die gesellschaftlichen Institutionen, in denen die Reflexion gerade über solche großen Fragen ihren Ort hat, kommen nicht mehr hinterher: Forschung und Wissenschaft reagieren nur noch auf die Revolutionen, die die Hohepriester des Silicon Valley wie am Fließband präsentieren – einem Fließband, das immer schneller läuft. Und weil kritische Stimmen fehlen, ist der Diskurs zu Big Data zwangsläufig von denen dominiert, die von Big Data am meisten profitieren: Neben den kalifornischen Hohepriestern sind das die selbsternannten Technologie-Evangelisten und Start-Up-Jünger auf der ganzen Welt. Oder die Marketing-Expertinnen, Data Miner, Berater, Finanzdienstleister, Marktforscherinnen und Lobbyisten. Flankiert von Informatikern, Mathematikerinnen und Systemarchitekten.

Sozial- und Politikwissenschaftler kratzen sich derweil am Kopf und sind erst noch dabei, die kritischen Fragen zu formulieren, die sie an das Phänomen Big Data haben, auch wenn sie dessen Vorteile schon fleißig nutzen. Geisteswissenschaftler scheinen völlig außen vor. International ist die Diskussion etwas weiter. Besonders Techniksoziologen, Kommunikations- und Kulturwissenschaftler im anglo-amerikanischen Ausland liefern wichtige Impulse. Im akademischen Bereich – vor allem

aber darüber hinaus. Der Erfolg ihrer Arbeit mag erste Hinweise geben, wie das Thema Big Data in der Wissenschaft besser reflektiert werden könnte.

Die Amerikanerin Danah Boyd etwa hat mit ihren „Six Provocations for Big Data“ schon 2011 einen ersten Überblick der großen Fragen zu Big Data formuliert. Boyd ist studierte Informatikerin, Medienwissenschaftlerin, sie äußert sich hauptsächlich als Techniksoziologin und ist mit ethnografischen Methoden vertraut. An der Harvard University und als Principal Researcher bei Microsoft arbeitet sie zudem ständig mit Kollegen aus verschiedensten Disziplinen zusammen, von Computerexperten über Softwareentwicklerinnen bis hin zu Kommunikationsdesignern, Ökonominen oder Policy-Experten. Sie publiziert in Fachjournalen, Blogs und der New York Times, präsentiert auf akademischen Konferenzen, bei Start-Up-Meetings und TED-Talks (einer internationalen Vortragsreihe zu aktuellen Themen, die zum Pflichtprogramm der digitalen Bohème gehört). Das ist gelebte Interdisziplinarität.

Viele ihrer Kollegen sind ähnlich stark fachübergreifend aufgestellt – in ihrer Vita, ihren Institutionen, ihren Quellen und ihren Zielgruppen. Es drängt sich der Verdacht auf, dass Interdisziplinarität für ihre Arbeit nicht nur einen Mehrwert liefert, sondern essenzielle Bedingung für sie ist. Der Grund dafür wäre naheliegend: Big Data ist verheißungsvoll, aber auch furchterregend, weil es so nah an den Kern des Menschen herankommt – so engmaschig viele Umstände seiner Existenz abzubilden vermag wie keine andere Technologie zuvor. So gesehen ist Big Data kein eigenes, neues Thema, sondern ein Prisma, durch das beinahe alle Aspekte unseres Lebens gebrochen werden. Manches erscheint so in völlig neuer und ungewohnter Weise.

Es ist die Zeit des Data-Turn in den Industriegesellschaften und dem Leben ihrer Bewohner. Der Umbruch betrifft auch viele traditionelle Gegenstände sozialwissenschaftlicher Forschung: soziale Ungleichheit (zum Beispiel: Wie wirkt sich Big Data auf Machtungleichgewichte aus?); Migration (Wie beeinflusst die extensive Social-Media-Kommunikation mit der Heimat die Integration in der Fremde?); Arbeitsmärkte (Wie verändert Big Data das Rekrutierungsverhalten von Arbeitgebern?); Sozialversicherungssysteme (Untergräbt Big Data im Versicherungswesen die Solidargemeinschaft?).

Big Data durchdringt immer weitere Teile des gesellschaftlichen Lebens, und es ändert sich ständig. Unter diesen Umständen auf eine adäquate oder gar abschließende Datenlage zu warten, ist aussichtslos. Der Input für die eigene Forschung muss zwangsläufig weiter ge-

fasst werden. Ganz oben stehen dabei vermutlich sogenannte APIs (Application Programming Interfaces). Das sind Zugänge, die es erlauben, aus verschiedensten Programmen, Datenbanken oder Anwendungen kontinuierlich (also in Echtzeit) Daten in strukturierter Form auszulesen. Im einfachsten Fall kann das heißen, einen Strom von Twitter-Nachrichten automatisiert auf bestimmte Begriffe zu untersuchen. Persönliche Erfahrung im Umgang mit und Offenheit gegenüber den Sozialen Medien ist für die produktive Nutzung dieser neuen Datenquellen natürlich von Vorteil. Vor allem aber verschafft es ein Verständnis für die beinahe grenzenlosen Möglichkeiten, menschliches Verhalten und gesellschaftliche Dynamiken abzubilden, zu analysieren. Und zu beeinflussen. Die erwähnten Big-Data-Forscher wenden sich direkt an eine viel größere Zielgruppe und beteiligen sich damit auch gleich an der öffentlichen Konversation zum Thema.

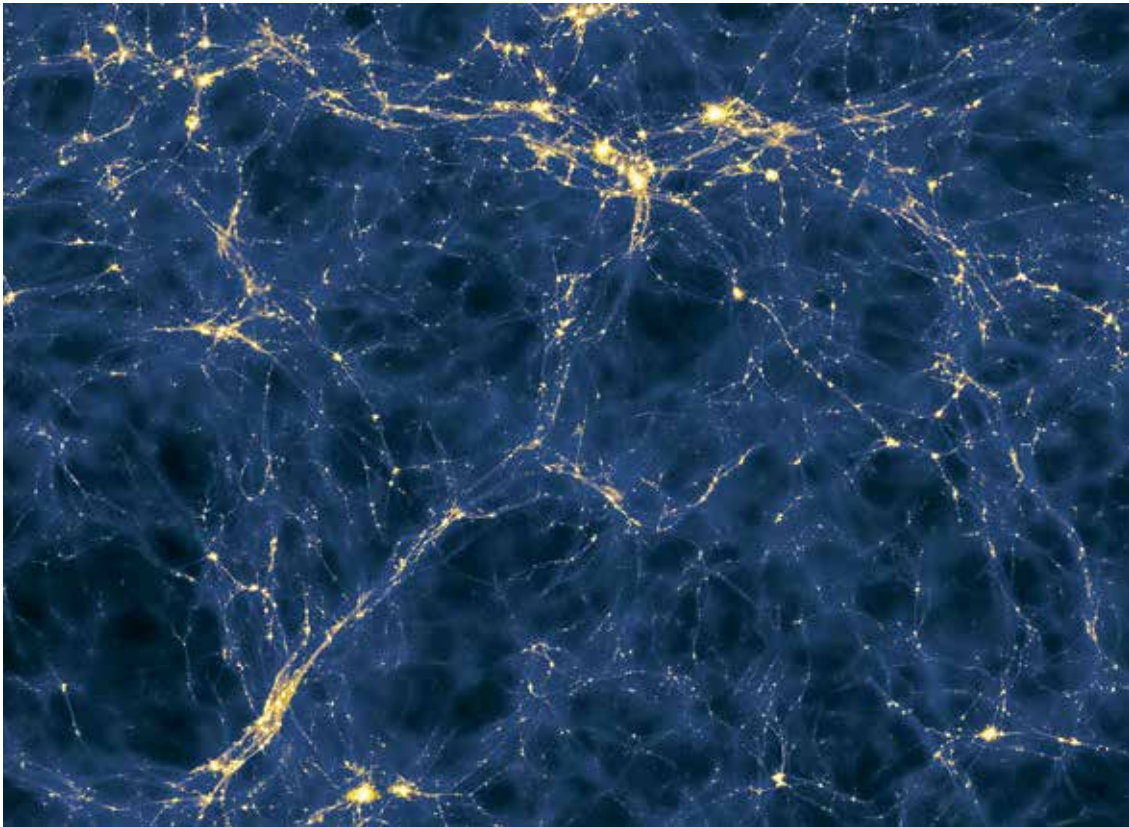
Das ist kein reiner Selbstzweck, sondern vor allem der Einsicht geschuldet, dass die zwangsläufig etwas langsamer mahlenden Mühlen der akademischen Wissensproduktion – Peer-Reviews, Edited Volumes, Monografien etc. – der atemberaubenden Massivität und Geschwindigkeit, mit der Big Data unser Leben beeinflusst und unsere Gesellschaft verändert, nicht mehr gerecht werden können. Ob die öffentliche Intervention oder der per Social Media vermittelte Schnappschuss der eigenen, möglicherweise unfertigen wissenschaftlichen Denkprozesse die einzigen oder einzig richtigen Wege sind, darüber kann man debattieren. Über neue Modi nachzudenken, wissenschaftliche Arbeit zu kommunizieren und zeitnah in öffentliche Debatten einzubringen, ist aber kein Verbrechen. Auch hier ist man außerhalb des Wissenschaftsstandorts Deutschland schon deutlich weiter.

Die Physiker mit dem Universum aus dem Rechner sind übrigens Waisenknaben (oder -mädchen) im Vergleich zu einigen ihrer Kollegen. Die vertreten nämlich die These, dass selbst unser „echtes“ Universum eine Simulation sei – wenn auch eine richtig komplizierte. Eine Freak-Idee? Tatsächlich geht diese steile Hypothese, die auch Philosophen, Mathematiker und Informatiker umtreibt, auf die alten Griechen zurück. Die Vorstellung, dass die Realität, inklusive unserer eigenen Existenz, durch ein kompliziertes Skript in einem unheimlich cleveren Automaten bestimmt wird, durch eine Art Software, die immer genau weiß, was wir fühlen, denken und tun – diese Vorstellung hat offenbar zu allen Zeiten ihre Anhänger gehabt. So ahnungslos, wie wir unser Leben gegenwärtig den Algorithmen ausliefern, steht Big Data kurz davor, diese kühne These auf unserem kleinen Planeten ein wenig Wirklichkeit werden zu lassen.

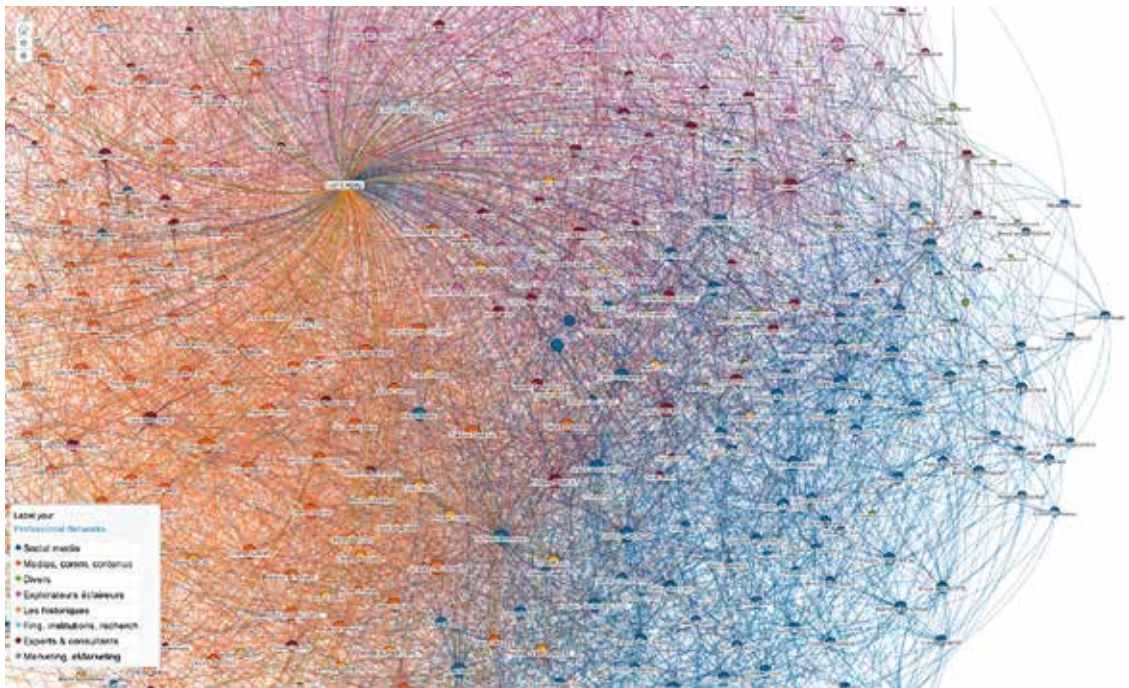


Der Sozialanthropologe und Physiker **Ranty Islam** war im März und April Fellow des Journalist-in-Residence-Programms am WZB. Bei der Deutschen Welle befasst er sich als Autor und Online-Editor vorwiegend mit Umwelt- und Auslandsberichterstattung. (Foto: privat)

Ranty Islam auf twitter: @2ranty



So sieht ein im Rechner simuliertes Universum aus.



Big Data heute: So sind Menschen in den Sozialen Netzwerken verbandelt.

Zweimal Datennebel: Das obere Bild zeigt ein im Rechner simuliertes Universum. Knoten, Fäden und Leerräume – so könnten die Vorläufer von Galaxien weit vor unserer Zeit ausgesehen haben. Das zweite Bild zeichnet am Beispiel eines Einzelnen und seiner Beziehungen nach, wie Menschen heute über soziale Netzwerke miteinander verbandelt sind. Die Abbildungen stellt Ranty Islam einander gegenüber. Der Social-Media-Redakteur bei der Deutschen Welle war im März und April journalist in residence am WZB. In seinem Beitrag für dieses Heft (S. 50f.) fordert er die Sozialwissenschaften heraus. Big Data dürfe nicht länger nur Werkzeug und Datenquelle sein, sondern müsse in seiner unwalzenden Rolle für die Gesellschaft diskutiert werden, schreibt der Physiker und Journalist.

[Abbildungen: Andrew Pontzen/Fabio Governato, UCL Mathematical and Physical Sciences, CC 2.0 BY; <https://www.flickr.com/photos/ucl-maps/15051460475/in/photostream/>. Luc Legay, CC 2.0 BY-SA; <https://www.flickr.com/photos/luc/5418037955/>]