



## Nicht Big sondern Smart Data ist entscheidend!

Yannick Haan und Patrick Rüter

Seit immer mehr Menschen in Deutschland bewusst wird, dass Daten der Rohstoff des 21. Jahrhunderts sind, ist auch hierzulande eine gesellschaftliche Debatte über Big Data ausgebrochen. Die Wirtschaft verbindet mit Big Data vor allem Effizienzsteigerung und die Möglichkeit, Kundenverhalten besser in Produktion und Vertrieb zu integrieren. Gleichzeitig werden mit der data driven economy auch völlig neue Geschäftsmodelle möglich, wie nicht zuletzt der Hype um die Industrie 4.0 und das Internet der Dinge zeigen. Für viele Bürger\_innen und Datenschützer\_innen hingegen ist Big Data wiederum nur ein letzter Schritt in eine vollüberwachte Gesellschaft, in der uns zunehmend der eigene Wille entzogen und durch Algorithmen ersetzt wird. Wie bei vielen Kontroversen muss man auch hier differenziert betrachten – Risiken und Chancen liegen nah bei einander.

### Von Small Data zu Big Data

Wenn man die radikale Entwicklung, die sich gerade durch den Einsatz von Big Data vollzieht, wirklich verstehen will, dann muss man sich unsere jetzige ‚Small Data‘ Welt anschauen. Auch heute versuchen wir bereits anhand von Daten und Zahlen die Welt zu verstehen und daraus Vorhersagen für die Zukunft zu treffen. Dies tun wir beispielsweise bei Wirtschaftsprognosen. Wir sammeln Daten, projizieren daraufhin eine Prognose für das Wirtschaftswachstum und planen anhand dieser Prognose Investitionen. Auch beim Wetter sammeln wir Daten und geben auf Basis dieser Daten eine Wetterprognose ab. Im Bereich der Wissenschaft versuchen wir bereits seit Jahrhunderten Daten zu sammeln, um

unsere Welt besser zu verstehen. Dieser menschliche Antrieb, mehr zu wissen, ist demnach weder neu noch revolutionär.

In den letzten Jahren hat sich dennoch ein qualitativer, vor allem quantitativer Prozess der Datennutzung vollzogen. Am Beispiel des Handels wird dies besonders deutlich. Früher haben wir uns ein Buch gekauft und es gelesen. Die einzigen Daten die dabei angefallen sind, war der Kauf des Buches selbst, das meist bei einer kleinen Buchhandlung gekauft wurde. Wir haben uns beim Kauf des Buches häufig auf die Empfehlungen der Buchhändlerin, unserer Freunde oder des Feuilletons verlassen. Als Amazon mit dem Betrieb angefangen hat, besaß der Onlinehändler daher auch noch eine eigene

Abteilung mit Menschen, die Bücher gelesen, bewertet und dann den Lesern empfohlen haben. Doch irgendwann kam Jeff Bezos, der CEO von Amazon, auf die Idee, die Daten, die er durch den Kauf eines Buches erhält, zu verknüpfen und zu nutzen, um den Kunden wiederum andere Bücher automatisiert zu empfehlen. Dank des elektronischen Kaufs und der Einführung von E-Books verfügen Amazon und andere Plattformen heute über immer mehr Daten und können diese nutzen, um individuelle Angebote des Unternehmens immer weiter an die Kunden anzupassen.

Dieser Wandel hat sich auch in anderen Branchen vollzogen. UPS konnte durch die Auswertung der Geodaten herausfinden, welche Strecken am besten zur Fahrt geeignet sind und wo die Staugefahr am höchsten ist. Was zunächst nach einer kleinen Optimierung klingt, hat dem Unternehmen laut eigenen Angaben 48 Millionen Kilometer und damit Benzin- und Fahrzeugkosten gespart, der Umwelt zudem 30.000 Tonnen CO<sub>2</sub>-Ausstoß erspart.<sup>1</sup> Der Einsatz von Big Data hat demnach nicht nur betriebswirtschaftlich wünschenswerte Folgen, sondern kann solche Innovationen auch gesellschaftlich nutzbar machen.

Der große Unterschied, der sich von Small Data zu Big Data vollzogen hat, liegt zunächst in der Datenmenge und der Fähigkeit, diese in Echtzeit auswerten zu können. Die Mobilmachung des Internets durch das Smartphone, das Nutzen sozialer Netzwerke und des Onlinehandels haben maßgeblich zur umfassenden Datenverfügbarkeit beigetragen. Der Preisverfall der Speichermedien hat diesen Prozess noch beschleunigt. Auch wenn die technische Entwicklung bereits weiter ist, fangen wir als Gesellschaften erst langsam an zu verstehen, wie wir mit großen Datenmengen umgehen können. Am Anfang einer Entwicklung, das sollte uns die Technikergangenheit gelehrt haben, ist es immer schwierig zu erkennen, in welcher Weise sich die Technik weiterentwickelt. Und doch lassen sich momentan drei große Entwicklungen erkennen.

## Aktuelle Entwicklungen

Erstens scheint es so, als ob große Unternehmen mit vielen Daten einen entscheidenden und von der Konkurrenz uneinholbaren wirtschaftlichen Vorsprung besitzen. Google, Facebook und Amazon besitzen so viele Daten über uns, dass sie uns genau kennen. Dieser Datenvorsprung führt im Internet zu einer ‚The winner takes it all‘ Dynamik. So kommt auf dem Suchmaschinenmarkt Google und dann ganz lange kein anderer Konkurrent. Im Online-Buchhandel Amazon und dann lange kein anderer Anbieter. Bei den sozialen Netzwerken ist Facebook ähnlich dominant. Man muss sich bei einer solch datengetriebenen Wirtschaft die ernste Frage stellen, ob neue Konkurrenten wirklich noch eine Chance gegen diese großen Unternehmen haben.

Eine zweite Entwicklung, die sich derzeit beobachten lässt, sind immer weitere Schritte hin zum Ende des Vergessens. Im Gegensatz zu unserem analogen, bleibt unser digitales Handeln oft für immer nachvollziehbar. Daten lassen sich in Sekundenschnelle und ohne Kosten vervielfachen. Die von vielen Datenschützern gewünschte Kontrolle über die eigenen Daten ist also technisch kaum umsetzbar. Das Ende des Vergessens liegt in der Struktur von Daten. Daran ändert auch das vor kurzem vom Europäischen Gerichtshof erlassene Urteil zum Recht auf Vergessen wenig. Es bleibt auch weiterhin die offene Frage welche Daten ‚vergessenswert‘ sind und welche Daten zum Nutzen der Gesellschaft eingesetzt werden können.

## Für die Bereitstellung von Daten bezahlen

Anstatt zu versuchen, ein technisch nicht umsetzbares Recht auf Vergessen zu etablieren, ist es umso dringender, sich politisch und gesellschaftlich mit Datenspeicherung und dem Ende des Vergessens auseinanderzusetzen. Denn die Datenerhebung selbst ist nicht das Problem, es ist die Intransparenz, wer, wo und zu welchem Zweck welche Daten erhebt.

<sup>1</sup> [http://www.pressroom.ups.com/pressroom/staticfiles/pdf/fact\\_sheets/UPS-big-data-infographic-900x2708.pdf](http://www.pressroom.ups.com/pressroom/staticfiles/pdf/fact_sheets/UPS-big-data-infographic-900x2708.pdf)

Wenn Geschäftsmodelle nur dann tragend sind, wenn sie auf genügend Kundeninformationen zurückgreifen können, dann sollten die Datenlieferanten (die Kunden) für ihren Faktoreinsatz auch entsprechend entlohnt werden. Oder zumindest die Möglichkeit bekommen, von diesem Tausch zurückzutreten. Andernfalls muss man zu dem Schluss kommen, dass es sich eben um keine nachhaltigen Geschäftsmodelle handelt. Mit entsprechenden Folgen für eine staatliche Förderung beispielsweise.

### **Die technische Individualisierung untergräbt den gesellschaftlichen Zusammenhalt**

Nicht zu unterschätzen sind außerdem die diskriminierenden Auswirkungen der individualisierten Datenauswertung. Durch die Individualisierung der Angebote durch Big Data als unrentabel erkannte Kund\_innen werden mit hoher Wahrscheinlichkeit von gewissen Angeboten, vom Vorzugspreis bis zum Kredit, ausgeschlossen. Das könnte und wird, wenn man es durchspielt genau die Menschen treffen, die ohnehin schon an den Rand der Gesellschaft gedrängt werden. Diese Entwicklung mag bei einfachen Handelsartikeln noch harmlos sein, entfaltet im gesellschaftlichen Trend jedoch ihr gefährliches Potenzial. Im Zeitalter der medialen Selbstoptimierung und digitalen Durchdringung der gesamten Lebenswelt (smart wearables, smart watches etc.) entsteht auch hier ein Macht- bzw. Meinungsmonopol, das Transparenz mit Demokratie verwechselt.<sup>2</sup> Das Vordringen der individualisierten Datenauswertung in das deutsche Versicherungswesen zeigt, wie schnell die absichernde Solidargemeinschaft plötzlich untergraben werden kann. So können sich beispielsweise Krankenkassenkunden durch Bonushefte oder Fitness-Apps eine Prämie durch gesunde Lebensführung erarbeiten. Dies entspricht faktisch einem verringerten Beitrag bzw. Finanzierungsanteil am Gesamtvolumen. Wer sich, aus welchen Gründen auch immer, nicht an solchen Programmen be-

teiligen kann oder will, erhält somit einen Nachteil. Denn die Kosten werden immer ungleicher verteilt. Auch die online-Nutzung von Nachrichten hängt zunehmend von den Vorlieben der jeweiligen Nutzer ab. Inhalte werden nicht mehr nach redaktionellem Ermessen, sondern nach Anzahl der Leser\_innen, eigenem Suchverhalten oder werbebezogenen und gefilterten Kriterien gelistet. Vielfältige Themen und Recherchen, die nicht in den Mainstream fallen, geraten so automatisch in den Hintergrund oder werden erst gar nicht angezeigt. So wird ein von der Mehrheit der Nutzer\_innen geprägtes (Nachrichten-) Bild erzeugt, das aber nicht deckungsgleich ist zu der gesamten Gesellschaft.

### **Auch die Datennutzung ist politisch**

Dabei ist gerade das klassische Versicherungsmodell ein gutes Beispiel, wie man sich politisch und gesellschaftlich bewusst gegen das individualisierte Optimum stellen kann. Obwohl wir über Daten zu den Erkrankungen und möglichen zukünftigen Erkrankungen von Menschen verfügen, haben wir uns politisch für ein System entschieden bei dem jede Bürger\_in, unabhängig ihrer aktuellen gesundheitlichen Verfassung, versichert wird. Wir haben uns, bislang zumindest, dafür entschieden die Risikodaten zu ignorieren und vielmehr gesellschaftliche Prinzipien zu formulieren. Wir können also durchaus auch politische Entscheidungen treffen, wie wir mit Daten umgehen und wie wir sie einsetzen wollen.

Die letzte und wohl wichtigste Konsequenz aus Big Data besteht darin, dass unser Leben ein Stück weit vorhersehbarer wird. Diese Entwicklung hat erst einmal viele positive Aspekte. Wenn wir es schaffen, die Zukunft ein wenig verlässlicher vorherzusagen, schaffen wir es auch Risiken besser zu berechnen. Wie bereits am Anfang erwähnt, liegt es in unserer Natur, Vorhersagen für unterschiedliche Bereiche zu treffen. Doch Big Data kann qualitativ nur so gut

<sup>2</sup> Vgl. PCAST: Big Data and Privacy. A technological perspective, 2014.

sein, wie die Informationen, auf denen sie beruht. Falsche Annahmen und Ausgangsdaten führen zu falschen Szenarien. Gerade die letzte Finanz- und Wirtschaftskrise, aber auch der Flash Crash<sup>3</sup> im Jahre 2010 haben eindrucksvoll bewiesen, wie fehlbar Wissenschaft und Technik sein können. Mit verheerenden Folgen für Menschen und Gesellschaften. Denn trotz aller verfügbaren Daten, wird gerade der oft irrational handelnde Mensch von Daten nicht berücksichtigt. Das Problem von Big Data ist folglich nicht Big Data oder die Vorhersage an sich, sondern unser gesellschaftlicher Umgang mit ihr.

### Smart Data und Smart Regulation

Wir müssen also nicht vor Technikgläubigkeit erstarren, denn selbst mit Big Data bleiben Vorhersagen nur Prognosen und können daher auch falsch sein. Wir brauchen allerdings die Transparenz bei Vorhersagen, auf welchen Annahmen sie beruhen und welche Daten hierzu genutzt werden. Denn diese Daten wiederum liefern uns auch die Möglichkeit der Wider-

legbarkeit von Handlungsvorhersagen. Es zeichnet sich bereits ab, dass nach Erhebung und Quantität der Daten, die qualitative Auswertung entscheidend ist für die Nutzbarmachung und Entscheidungsfindung. Dies eröffnet auch neuen Marktteilnehmern die Chance, ihre Dienste anzubieten. Neue Tätigkeitsprofile wie data harvesting und data mining zeigen schon dem Namen nach, dass die Claims noch nicht endgültig vergeben sind. Große Datenmengen (Big Data) und intelligente Daten (Smart Data) bleiben immer noch verschiedene Dinge.

Was wir jetzt brauchen, ist eine Rückbesinnung auf unsere Grundrechte und Grundwerte, die mit der Digitalisierung nur das Medium gewechselt haben. Das Ernten und Nutzen von Daten kann genauso reguliert werden wie andere Rohstoffe auch. Wenngleich Big Data die Position der Oligopole gestärkt hat, es wären nicht die ersten Fälle, in denen ein kluges Kartellrecht nicht auch angewandt werden könnte. Rechtsstaatlichkeit und Ordnungspolitik sind auch und vor allem in der Wirtschaft hohe Güter.

---

<sup>3</sup> Im Mai 2010 stürzten die US-amerikanischen Börsenkurse plötzlich rasant ab, erholten sich nach kurzer Zeit jedoch wieder. Der computergetriebene und auf Algorithmen beruhende Hochfrequenzhandel gilt als einer der Ursachen für diese Kettenreaktion.

#### Über die Autoren

**Yannick Haan** ist Online-Berater, Netzaktivist und Co-Autor des Buchs „Gesellschaft im digitalen Wandel“, erschienen bei der Friedrich-Ebert-Stiftung.

**Patrick Rüter** ist Referent für den Managerkreis der Friedrich-Ebert-Stiftung und dort unter anderem für den Bereich Digitale Wirtschaft und Gesellschaft zuständig.