

Big Data, Small Money, No Privacy?

Was bekommt der Verbraucher für seine persönlichen Daten und wie viel ist Datenschutz wert?



Sören Preibusch,
Post-Doctoral
Researcher, Micro-
soft Research,
Cambridge, UK
spr@microsoft.
com

Lohnt sich der Verzicht auf Privatsphäre? Wie sieht Fairtrade bei Big Data aus, so dass der Nutzer nicht nur Produkt, sondern auch Profiteur wird?

Die Werthaltigkeit von Big Data ist unbestritten und schürt die mit dem Trendwort verknüpften Hoffnungen. Die Verarbeitung grosser Datenmengen ist und bleibt Treiber der Informationstechnologie. Neu ist, dass die Daten der fortlaufenden Aufzeichnung unseres menschlichen Handelns und Miteinanders entstammen. Der Arbeitskreis Big Data beim Bundesverband Informationswirtschaft, Telekommunikation und neue Medien (BITKOM) sieht den wirtschaftlich nutzbaren Erkenntnisgewinn als definierendes Merkmal von Big Data¹ und zählt als typische Beispiele unter anderem Verbindungsdaten in der Telekommunikation, Zugriffsstatistiken auf Webseiten und Verbrauchsdaten von Strom und Wasser auf – kurzum das Protokoll von Konsumententscheidungen und Nutzerverhalten.

Anschaulich zeigt sich der Wert aggregierter Nutzerdaten im Auf und Ab der Bewertungen von Plattformen wie AOL und MySpace: Als die Nutzer gingen, fiel der Unternehmenswert. Ein weiteres Beispiel: Der Link-Sammeldienst Digg wurde 2012 für einen einstelligen Millionen-Dollar-Betrag gekauft. Vier Jahre zuvor, bevor die Nutzer zu Alternativen wie Reddit abwanderten, lag die Bewertung noch um zwei Grössenordnungen höher².

Im Oktober 2012 kündigte der spanische Telekommunikationskonzern Telefónica an, Standortdaten seiner Kunden aufbereiten und verkaufen zu wollen, mit dem Ziel, Wert aus Big Data zu schaffen³. Drei Wochen später zog Telefónica das Vorhaben für Deutschland zurück, mit Verweis auf «Datenschutz und Kundenzufriedenheit»⁴. Jenseits der rechtlichen und ethischen Beurteilung verdeutlicht dieses vorerst gescheiterte Big-Data-Vorhaben das öf-

fentliche Unbehagen mit der kommerziellen Nutzung grosser Mengen personenbezogener Daten. Dieses Unbehagen speist sich aus zwei Wahrnehmungen: einerseits dem Verlust an Privatsphäre, andererseits der einseitigen Verteilung der Datenrente zugunsten des Dienstanbieters.

Die wissenschaftliche Datenlage zu diesen Faktoren ist zurzeit noch wenig ergiebig, konnte aber bereits erste Befunde zu den Einstellungen von Konsumenten liefern. In einem Feldexperiment mit 168 Teilnehmern fanden CARRASCAL ET AL. (2011), dass Verbraucher zwar bereit sind, ihre persönlichen Daten für Zwecke der Dienstverbesserung bereitzustellen, einer Versilberung derselben Daten durch denselben Anbieter aber durchwegs wenig aufgeschlossen gegenüberstehen, auch wenn sie darüber Bescheid wissen. Es scheint, dass Online-Nutzer es nur widerwillig hinnehmen, nicht an der Datenrente, die Unternehmen schöpfen, beteiligt zu sein.

Der Vorstoss Frankreichs, eine Steuer auf das Halten persönlicher Daten zu erheben, verspricht zumindest ansatzweise eine Vergemeinschaftung der Gewinne aus Datenressourcen: Unternehmen, die grössere Mengen an Nutzerdaten halten und daraus Gewinn schöpfen, sollen einen Teil davon an den Nationalstaat abführen⁵. Zwar würde der bisher kostenlose Datenbeitrag jedes einzelnen Verbrauchers nicht in Euro und Cent abgegolten, wohl aber über den Fiskus an die Nutzergemeinschaft insgesamt zurückgeführt. Es muss gleichwohl erörtert werden, ob nicht der Verbraucher bereits heute geldwerte Vorteile aus datenfinanzierten Diensten schöpft. Für Grossbritannien zum Beispiel geht eine Boston Consulting Group-Studie davon aus, dass Nutzer von kostenlosen werbefinanzierten Diensten pro Jahr 5 Milliarden USD Gegenwert erhalten, was dem Doppelten der entrichteten Internetzugangsentgelte entspräche⁶.

Hier schliesst sich der Kreis zur oben angesprochenen Wahrnehmung und Tatsache, dass Big Data zulasten der Privatsphäre geht⁷. Kurzum: Sind Nutzen und Gewinn aus Big Data dem

Online-Nutzer genügend wert, um dafür auf Aspekte seiner Privatsphäre zu verzichten?

Entstehung und Erschliessung von Wert in Big Data

Es ist unbestritten: Unternehmen profitieren vom Errichten und Ausbeuten nutzerbezogener Datenberge. Grundlage dafür sind Datenschutzvereinbarungen und Dienstnutzungsbestimmungen, mit denen die Betreiber weitreichende Einverständniserklärungen vom Nutzer hinsichtlich Erhebung, Speicherung und Nutzung von Daten einholen. Die erhobenen Daten durchlaufen mehrere Verarbeitungsschritte, damit aus Big Data schliesslich gewinnbringende Erkenntnis wird.

Der Wert, der in Big Data steckt, muss erschlossen werden. Ein grosser Datenbestand allein nützt nichts, sondern verursacht zunächst Kosten für Speicherung und Langzeiterhalt. Selbst bei automatisiert protokollierten Daten wie Verbrauchsprotokollen entsteht Aufwand für Fehlerkontrolle, Datenbereinigung und -anreicherung und Deduplizierung. Der Informationsdienstleister Experian weist in seiner Bilanz für 2012 «data and information technology costs» in Höhe von 400 Millionen USD aus⁸.

Anschliessend bedienen sich die Dienstleister der Methoden des Data Mining, um nicht-triviales Wissen aus den aufbereiteten Daten zu ziehen. Dazu bedarf es der Methodenkenntnis, aber ganz praktisch auch notwendiger Ressourcen wie Speicherplatz und Rechenleistung. So kostet beim Cloud-Anbieter Windows Azure ein mittelgrosser Rechner mit 10 Terabyte lokalem Speicher und einer 50 Gigabyte grossen Datenbank knapp 900 US-Dollar pro Monat⁹. Zum Vergleich: Der Online-Marktplatz eBay hält 9 Petabyte an Daten vor¹⁰.

Fortgeschrittene Methoden des Big Data sind also erstens häufig zu teuer, um vom einzelnen Verbraucher in Heimarbeit durchgeführt werden zu können. Zweitens fehlt dem einzelnen Datensubjekt die Gesamtschau auf den Datenbestand. Eine Verringerung der Datenbasis schmälert den möglichen Erkenntnisgewinn überproportional.

Ein anschauliches Beispiel sind Staumeldungen. Die Google-Navigation berücksichtigt für die Routenplanung die aktuelle Verkehrslage – zweifelsohne zum Vorteil des Nutzers. Das Wissen um die Verkehrsstockung zieht der Betreiber aus aggregierten Zeit- und Standortdaten von Android-Smartphones¹¹. Aus lediglich seinen eigenen Daten könnte der Nutzer den Stau erst erkennen, wenn er bereits darin stünde. Auch Produktempfehlungen im Onlinehandel speisen sich aus den Bestellhistorien der

vielen anderen Kunden. Collaborative filtering¹² findet Produkte, die Kunden mit ähnlichen Interessen bestellt haben. Ohne Wissen um die Bestellungen aller Kunden müssten Empfehlungen auf Produktähnlichkeit basieren.

Schliesslich sei noch auf die wichtige Rolle von Intermediären hingewiesen. Sie vermitteln zwischen Dienstleister und Kunden und ist typisch für die re-konfigurierten Wertschöpfungsketten der digitalen Ökonomie. Für den Intermediär bietet Big Data die Chance, über die Vermittlerrolle hinauszuwachsen. Beispiele

Sind der Nutzen und der Gewinn aus Big Data dem Online-Nutzer genügend Wert, um dafür auf Aspekte seiner Privatsphäre zu verzichten?

sind Zahlungsabwickler, die ihre Datenbestände zur Betrugsvermeidung nutzen, und Telekommunikationsanbieter, die mehr als nur Daten durchleiten wollen (wie Telefónica). Auch Facebook vermittelt zwischen seinen Mitgliedern.

Zusammenfassend bleibt festzuhalten: Der Verbraucher kann den Wert, der in Big Data steckt, nicht allein erschliessen. Er selbst hat nur «Small Data», und deren Wert ist unterproportional niedrig. Es bedarf der Zusammenarbeit von Dienstleister und Dienstinhaber¹³.

Geld gegen Daten: der Wert der Privatsphäre

Was also ist eine angemessene Kompensation des Verbrauchers für die Datenpreisgabe? Wie oben angedeutet, muss der Gegenwert nicht bar ausgezahlt werden – eine Beteiligung des Kunden am Wert von Big Data kann auch durch höhere Servicequalität umgesetzt werden. Es gilt jedoch: Der Verbraucher tauscht Daten gegen Wert. Worauf müsste der Verbraucher verzichten, wenn er sich gegen die Datenpreisgabe entscheidet? Was ist der Wert der

Kurz & bündig

Der Verbraucher kann den Wert, der in Big Data steckt, nicht allein erschliessen. Er selbst hat nur «Small Data», und deren Wert ist unterproportional niedrig. Es bedarf also einer Zusammenarbeit von Dienstleister und -nachfrager. Eine angemessene Beteiligung des Nutzers an der Datenrente kann monetär oder nicht-monetär erfolgen (z.B. durch höhere Dienstqualität). Idealerweise wird die Wahl dem Verbraucher selbst überlassen. Eine Entlohnung für Datenpreisgabe entspricht in der Umkehrung einem Preisaufschlag für mehr Privatsphäre. Zahlungsbereitschaften für besseren Datenschutz lassen sich experimentell ermitteln. In einer Feldstudie war ein Drittel der Teilnehmer bereit, 1 Euro zu entrichten, um nicht ihre Mobilrufnummer angeben zu müssen. Nur etwa jeder zehnte zahlte einen Obolus, um keine Werbung per E-Mail zu erhalten.

gewonnenen Privatsphäre, wenn man seine Daten nicht weitergibt?

Welchen Wert Verbraucher bestimmten Produkten beimessen, ist schwer zu ermitteln, und noch schwieriger wird es bei immateriellen Versprechen wie «mehr Privatsphäre». Der naive Ansatz, Verbraucher in einer Umfrage zu bitten, ihre Wertschätzung für etwas zu nennen, läuft ins Leere. Selbst wenn man sich zu einem Euro-Betrag durchringt, scheint es lohenswert, einen höheren Betrag anzugeben, wenn man entlohnt werden möchte, bzw. einen niedrigeren Betrag anzugeben, wenn man befürchtet, zur Kasse gebeten zu werden.

Zahlungsbereitschaften lassen sich experimentell ermitteln und liefern eine Untergrenze für den beigemessenen Wert. Solange Konsumentscheidungen freiwillig gefällt werden, wird man nur so viel Geld für etwas ausgeben, wie es Nutzen stiftet. Jemandem, der eine Flasche Cola für 3 Euro kauft, war die Erfrischung zu dem Zeitpunkt mindestens diesen Betrag wert. Wer den Preis dagegen als zu hoch ansieht und auf den Konsum verzichtet, hat eine geringere Zahlungsbereitschaft und also eine geringere Wert-

Der Verbraucher kann den Wert, der in Big Data steckt, nicht allein erschliessen. Er selbst hat nur «Small Data», und deren Wert ist unterproportional niedrig.

schätzung. Für jeden Preispunkt lässt sich so der Anteil der Bevölkerung mit einer bestimmten Zahlungsbereitschaft ermitteln – durch Beobachtung von realen Kaufentscheidungen oder experimentell in Labor- und Feldstudien. Der funktionale Zusammenhang zwischen Preis und Nachfrage ist dabei aussagekräftiger als eine gemittelte Zahlungsbereitschaft.

Dasselbe Kalkül (Wertschätzung \leq Zahlungsbereitschaft) lässt sich auf Entscheidungen für oder wider die Datenpreisgabe anwenden. Das betrifft einerseits Preise für Datenschutztechnologien (tatsächlich z.B. 50 Euro für einen Blickschutzfilter für den Bildschirm), andererseits Preisaufschläge, die Verbraucher bereit sind, für eine Produkt- oder Dienstvariante mit besserer Privatsphäre zu entrichten (fiktiv zum Beispiel 5 Euro mehr im Monat für

einen Handyvertrag ohne langfristige Speicherung von Standortdaten).

Experimente, die Zahlungsbereitschaften für Datenschutz in einer konkreten Konsumsituation messen, haben gegenüber Umfragen neben höherer Zuverlässigkeit den weiteren Vorteil des Kontextes. Die Bereitschaft, persönliche Daten anzugeben, lässt sich nicht loslösen vom Datenempfänger und das Vertrauen in ihn, die situative Gestaltung der Datenabfrage (das sog. Framing) und Eigenschaften der zugrunde liegenden Transaktion (Produkteigenschaften, Dringlichkeit usw.).

Gemessene Zahlungsbereitschaften für Privatsphäre

Zwar ist die empirische Evidenz zur Zahlungsbereitschaft für Privatsphäre noch spärlich, doch die Anzahl an experimentellen Arbeiten zum Thema wächst. Belastbare Daten lieferte zuletzt ein grosses Feld- und Laborexperiment für die European Network and Information Security Agency (ENISA). Mehr als 500 Teilnehmer im Labor und viele weitere hundert Webnutzer in der Feldstudie waren eingeladen, bis zu zwei Kinokarten zu kaufen. Dabei hatten sie die Wahl zwischen einem datenschutz-freundlichen und einem datenschutz-unfreundlichen Anbieter. Letzterer verlangte je nach Experimentdesign zusätzlich die Mobilrufnummer des Käufers beziehungsweise dessen Einverständnis, die E-Mail-Adresse auch für Werbezwecke nutzen zu können. Im Gegenzug dafür bot dieser Anbieter die Kinokarten in einigen Experimentalbedingungen um einen halben Euro günstiger an.

Wenn kein Preisunterschied bestand, bevorzugten die Studienteilnehmer mehrheitlich den datenschutz-freundlichen Anbieter; dieser erzielte einen Marktanteil von 83% bzw. 62%.

Etwa ein Drittel der Probanden war bereit, einen Euro mehr zu bezahlen, um nicht ihre Mobilrufnummer angeben zu müssen¹⁴. Um keine Werbe-E-Mails zu erhalten, öffnete immerhin noch rund jeder Zehnte sein Portemonnaie.

Ist also ein Mehr an Privatsphäre einen Euro wert? Zumindest für den Anwendungsfall Electronic Commerce sind die relativen Zahlungsbereitschaften für geringere Datenerhebung und -nutzung plausibel generalisierbar. Für andere Szenarien und Anwendungsgebiete von Big Data sind im Zweifel weitere Studien durchzuführen. Die Kontextabhängigkeit der Datenschutz-Präferenzen, einschliesslich der wahrgenommenen Relevanz von Daten, könnte z.B. zu einer verminderten Bereitschaft führen, die Telefonnummer für die Wettervorhersage anzugeben.

Literatur

- JUAN PABLO CARRASCAL ET AL., Your browsing behavior for a Big Mac: Economics of Personal Information Online, 2011.
- NICOLA JENTZSCH ET AL., Study on monetising privacy. An economic model for pricing personal information, ENISA, 2012.
- SÖREN PREIBUSCH ET AL., The privacy economics of voluntary over-disclosure in Web forms, 11th Workshop on the Economics of Information Security, 2012.

Zu diskutieren bleibt, ob ein Euro eher viel oder eher wenig Geld für extra Privatsphäre ist. Hier bieten sich drei Vergleichsmaßstäbe an: erstens das Einkommen des Kunden, zweitens der Preisaufschlag relativ zum Produktpreis bzw. drittens relativ zu den Rabatten, die Kundenkarten versprechen. Selbst für Studierende mit existenzsicherndem Stipendium («BAföG») oder für Bezieher der sozialen Grundsicherung («Hartz IV») liegt der Aufschlag bei deutlich unter 1% des monatlichen Bareinkommens. Im Vergleich zum Ticketpreis entspricht die Privatsphäre-Gebühr 13% bis 18% des Ticketpreises bei Kinokarten. Das ist zwar ein deutlicher Obolus, aber immer noch günstiger als die kleinste Packung Popcorn. Am aufschlussreichsten ist der Vergleich mit Kundenkarten, die Vergünstigungen beim Einkauf versprechen und als Datenschutzrisiko verschrien sind: Kartennutzer geben bei Vertragsschluss ihre persönlichen Details an (einschliesslich Telefonnummer) und fortan Details zu jedem ange-rechneten Einkauf. Sofern die geldwerten Ver-günstigungen überhaupt ermittelt werden kön-nen, entsprechen sie maximal 5% Rabatt (Payback) und sind damit trotz eines umfang-reicheren Pakets an persönlichen Informatio-nen niedriger¹⁵.

Fazit

Daten sind Schmiermittel und Werkstoff des digitalen Wirtschaftens. Durch individualisier-ten Konsum erlangen viele dieser Daten Perso-nenbezug. Durch Datenaufbereitung und -an-reicherung extrahieren Unternehmen Wert aus Big Data. Für den einzelnen Verbraucher ist

diese Erschliessung typischerweise nicht mach-bar. Auch wenn Big Data bereits jetzt das Ver-sprechen auf bessere Produktqualität birgt, wünschen sich viele Verbraucher eine faire Beteiligung am Wert ihrer Daten. Eine Entschä-digung für Datenpreisgabe entspricht einem Preisaufschlag für Datengeheimhaltung.

Der Wert von Privatsphäre ist schwer mess-bar und stark kontextabhängig. Als Anhalts-punkt dienen Preisaufschläge, die Verbraucher bereit sind, für besseren Datenschutz zu zah-

Der Wert von Privatsphäre ist schwer messbar und stark kontextabhängig. Als Anhaltspunkt dienen Preisaufschläge, die Verbraucher bereit sind, für besseren Datenschutz zu zahlen.

len. Zuletzt zeigte eine gross angelegte Studie, dass ein Drittel der Online-Shopper bereit ist, einen Euro mehr zu zahlen, um nicht ihre Mo-bilrufnummer angeben zu müssen. Gleichzeitig belegen andere Feldexperimente wie PREIBUSCH ET AL. (2012), dass Verbraucher weiterhin bereit sind, ihre Daten freiwillig und umsonst bereit-zustellen. Die Aussicht auf bessere Dienstquali-tät kann ein starker Treiber sein, aber auch Hilfsbereitschaft oder die Persönlichkeit.

Wichtig sind Transparenz, ein klares und eingelöstes Nutzenversprechen für den Ver-braucher und eine faire Beteiligung an der Datenrente. Ob diese besser monetär oder nicht-monetär erfolgt, ist im Einzelfall zu ent-scheiden. Idealerweise wird die Wahl dem Ver-braucher selbst überlassen. ■

Fussnoten

¹ <http://www.bitkom.org/de/wir_ueber_uns/70822.aspx>.

² <<http://bits.blogs.nytimes.com/2012/07/12/betaworks-buys-whats-left-of-social-news-site-digg/>>.

³ «establishment of Telefónica Dynamic Insights, a new global business unit, dedicated to identifying and unlocking the potential opportunities for creating value from «big data.»» (<http://pressoffice.telefonica.com/documentos/nprensa/Telefonica_Dynamic_Insights_press_release_finat.pdf>).

⁴ <<http://www.heise.de/newsticker/meldung/Telefonica-Keine-Analyse-von-Bewegungsdaten-in-Deutschland-1741717.html>>.

⁵ <http://www.redressement-productif.gouv.fr/files/rapport-fiscale-lite-du-numerique_2013.pdf>.

⁶ The Boston Consulting Group, The Digital Manifesto, How Companies and Countries Can Win in the Digital Economy, Januar 2012.

⁷ <<http://www.danah.org/papers/talks/2010/WWW2010.html>>.

⁸ <<http://www.experianplc.com/~media/Files/E/Experian-V2/pdf/investor/reports/2012/experian-ar-2012.pdf>>.

⁹ <<http://www.windowsazure.com/de-de/pricing/calculator/?scenario=full>>.

¹⁰ <http://www.computerworld.com/s/article/print/9224180/What_the_big_deal_about_Hadoop?taxonomyName=Business+Intelligence%2FAnalytics&taxonomyId=9>.

¹¹ <<http://www.pcwelt.de/ratgeber/Stau-Warnung-So-erkennt-Google-Maps-die-Verkehrslage-373385.html>>.

¹² Kollaboratives Filtern ist ein Data-Mining-Verfahren, das sich auf Ähnlichkeiten im Nutzerverhalten stützt, z.B. das Kaufen oder Bewerten bestimmter Artikel. Für die Vorhersage, ob ein bestimmter Nutzer einen Artikel kaufen wird, wird das Kaufverhalten von anderen Nutzern herangezogen, die mit dem aktuellen Nutzer in der Vergangenheit für andere Artikel ähnliches Verhalten aufwiesen.

¹³ Die Informatik forscht an dezentralen Architekturen mit verteilter Datenhaltung.

¹⁴ Der überwiegende Teil der Probanden kaufte zwei Kinokarten. Pro Ticket war der datenschutz-freundliche Anbieter 0,50 Euro teurer. In der Summe wurde also 1 Euro für mehr Privatsphäre bezahlt. Eine Falschangabe der Mobilrufnummer war ausgeschlossen: Ein Bestätigungscode wurde per SMS zugestellt.

¹⁵ <<http://www.test.de/Kundenkarten-Verfuehrung-a-la-carte-1678989-1678901/>>.