# Ethische Herausforderungen für Unternehmen im Umgang mit Big Data

satw it's all about technology

# Inhaltsverzeichnis




**Impressum**

Autoren: Christian Hauser, HTW Chur (Leitung), Helene Blumer, HTW Chur, Markus Christen, Universität Zürich, Lorenz Hilty, Universität Zürich/Empa, Markus Huppenbauer, Universität Zürich und Tony Kaiser, SATW

Redaktion: Beatrice Huber

Chefredaktor: Urs von Stockar

Titelbild: Steffon Davis at steffondavis@gmail.com

1. Auflage, Januar 2017

# Einleitung

Seit einigen Jahren steht das Stichwort «Big Data» stellvertretend für einschneidende Veränderungen, welche die zunehmende Durchdringung der Gesellschaft mit digitaler Technologie mit sich bringt. Eine präzise Definition für diesen jungen Begriff gibt es zwar noch nicht. Er verdeutlicht aber, dass heutzutage auf einfache Weise grosse Mengen sehr unterschiedlicher Daten erfasst, gespeichert und analysiert werden können. Nicht nur Computer, Smartphones und tragbare Sensoren, sondern auch Autos, Haushaltsgeräte und Gebäude erfassen dank eingebetteter Informations- und Kommunikationstechnologie (IKT) routinemässig, wo wir sind, was wir tun und mit wem wir kommunizieren. Dies geschieht nicht unbedingt in der Absicht, Menschen gezielt zu überwachen. Vielmehr ist es eine inhärente Eigenschaft digitaler Technologie, Daten zu erzeugen: «Wüsste» beispielsweise ein Mobilfunknetz nicht, wo sich ein Smartphone befindet, wäre dieses nicht erreichbar. Für viele Unternehmen erscheint es heute undenkbar, Entwicklung, Produktion und Verkauf ohne massgebliche Unterstützung von IKT umzusetzen. Dies erzeugt automatisch Daten über Prozesse, die früher quasi im Verborgenen abliefen – man vergleiche etwa das Stöbern eines Konsumenten in einem stationären Buchladen mit der Suche im virtuellen Buchladen. Die so erzeugten Daten können zudem auf zunehmend einfachere und billigere Weise gespeichert und auf immer raffiniertere Weise ausgewertet werden. Damit werden Prozesse von Produktion und Konsum und damit letztlich die einzelnen Menschen auf eine Weise erfassbar, die vor wenigen Jahren noch undenkbar erschien.

Tiefgreifende Entwicklungen wecken Hoffnungen und Ängste: Manche sehen in Big Data das «Öl des 21. Jahrhunderts» und in den Daten eine enorme Ressource für Innovation. Andere halten Big Data für eine fundamentale Bedrohung für Freiheit und Privatsphäre, ein dämonisches Instrument in einem Orwell'schen Überwachungsstaat. Beide Szenarien mögen überzeichnet sein – sie verdeutlichen aber, dass mit Big Data schwierige ethische Fragen verbunden sind: Was bedeuten Grundwerte wie Selbstbestimmung, Solidarität und Privatsphäre in einer Big-Data-Welt? Stellen gewisse Daten ein öffentliches Gut dar? Sind wir verpflichtet, manche Daten preiszugeben, um die Gesellschaft effizienter oder sicherer zu machen?

Diese Broschüre geht solchen Fragen für eine spezifische Einsatzform von Big Data nach, die jeden betreffen: die Nutzung von Daten durch Unternehmen an der Schnittstelle zu ihren Kunden. Personalisierte Werbung, massgeschneiderte Angebote oder individualisierte Preisgestaltung sind Beispiele solcher Anwendungen. Welche ethischen Fragen werden dadurch aufgeworfen? Welche Werte müssen in solchen Fällen gegeneinander abgewogen werden? Was sind realistische Chancen und Risiken von Big-Data-Anwendungen im Consumer-Bereich? Solche Fragen können zwar nicht endgültig beantwortet, aber zumindest in strukturierter Form vorgestellt und diskutiert werden. Nach methodischen Vorbemerkungen werden dazu die begrifflichen und rechtlichen Grundlagen dargestellt sowie die wissenschaftliche Literatur und Medienberichte zu Big Data ausgewertet. Danach illustrieren fünf Beispiele aktuelle und künftig mögliche Big-Data-Anwendungen durch Unternehmen und die damit verbundene Nutzung von Kundendaten. Anhand dieser Beispiele werden die ethischen Herausforderungen von Big-Data-Anwendungen verdeutlicht. Der Bericht schliesst mit Handlungsempfehlungen für Politik und Unternehmen.

# Vorbemerkungen

## Methodik

Ziel dieses Berichts ist es, aktuelle und zukünftige Big-Data-Anwendungen im Bereich Wirtschaft sowie die ethischen Werte, die von diesen Anwendungen berührt werden, zu identifizieren und zu diskutieren. Im Fokus steht die Schnittstelle zwischen Unternehmen und ihren Kunden. Die Ausführungen stützen sich auf qualitative und quantitative Literaturanalysen, Experteninterviews und Workshops.

Um eine Einschätzung der Literatur zum Themenbereich Big Data zu erhalten, wurde eine bibliometrische Untersuchung in zwei wissenschaftlichen Literaturdatenbanken (Web of Science und Scopus)[1] sowie in der Mediendatenbank Factiva[2] durchgeführt. Die identifizierten Publikationen dienten für eine Ausdifferenzierung der in der Literatur diskutierten ethischen Aspekte zu Big Data. Daraus ergab sich ein Stichwortset zur Charakterisierung von Publikationen, in denen eine ethische Begrifflichkeit in Titel, Abstract oder Keywords vorkommt. Hochzitierte Publikationen wurden auch qualitativ ausgewertet. Die Suche erfolgte zunächst explorativ über den gesamten verfügbaren Zeitraum; für die Analyse wurde die Suche auf die Jahre 2006 bis 2015 eingeschränkt.

Experten wurden auf zwei Arten eingebunden. Im Rahmen von zwei Workshops identifizierten insgesamt 22 teilnehmende Fachpersonen bestehende und künftig mögliche Big-Data-Anwendungen durch Unternehmen und diskutierten die damit assoziierten ethischen Fragen bei der Nutzung von Kundendaten. Ergänzend wurden zwei Interviews mit Fachleuten durchgeführt. Die Experten kamen aus den Branchen Banken und Versicherungen, Unternehmensberatung, Marketingdienstleister, Soft- und Hardwarehersteller, Detailhändler sowie Telekommunikations- und Transportunternehmen. Darüber hinaus wirkten kantonale und eidgenössische Datenschutzbeauftragte sowie Vertreter der Wissenschaft mit.

Die beteiligten Fachleute identifizierten und diskutierten fünf beispielhafte Tätigkeitsfelder, in denen Unternehmen Big-Data-Anwendungen aktuell nutzen bzw. zukünftig stärker nutzen werden:

- Debitorenausfälle verhindern
- Risikomanagement verbessern
- Angebotskonditionen massschneidern
- Effizienz der Werbemassnahmen steigern
- Innovationen hervorbringen und neue Umsatzquellen erschliessen

Im Kapitel «Big-Data-Anwendungen» werden diese Tätigkeitsfelder anhand von aktuellen und sich zukünftig abzeichnenden Praxisanwendungen illustriert.

Ausgehend von diesen Tätigkeitsfeldern identifizierten und diskutierten die Experten acht ethische Normen und Werte, die von Big-Data-Anwendungen berührt werden:

1. Schutz der Privatsphäre
2. Gleichheit und Nichtdiskriminierung
3. Informationelle Selbstbestimmung
4. Kontrolle der eigenen (digitalen) Identität
5. Transparenz
6. Solidarität
7. Kontextuelle Integrität
8. Eigentums- und Urheberrecht

Basierend auf diesen Normen und Werten werden im Kapitel «Ethische Debatte» die beschriebenen Big-Data-Anwendungen einer ethischen Analyse unterzogen. Aus diesen Erkenntnissen werden im Folgekapitel Handlungsempfehlungen für Politik und Unternehmen abgeleitet.[3]

---

[1] Web of Science (WoS): **https://apps.webofknowledge.com**; Scopus: **http://www.scopus.com**. Das Suchstichwort war in allen Datenbanken der Ausdruck «big data» (WoS: unter «topics»; Scopus: in der Kategorie «title, abstract, keywords»).

[2] Die von Bloomberg unterhaltene Datenbank Factiva enthält die wichtigsten internationalen Printmedien wie auch Beiträge zahlreicher anderer Informationskanäle vorwiegend aus dem wirtschaftlichen Bereich: **https://global.factiva.com**.

[3] Aus Gründen der Lesbarkeit wird auf Quellenverweise im Text generell verzichtet. Stattdessen findet sich am Schluss der Broschüre eine Auswahl an weiterführender, vor allem deutschsprachiger Literatur.

# Begriffliche und rechtliche Grundlagen

Der Begriff «Big Data» findet sich erst in jüngster Zeit in der Literatur. Eine einheitliche Definition existiert zwar bislang nicht. Eine systematische Untersuchung der Begriffsverwendung durch Jonathan Stuart Ward und Adam Barker (siehe dazu empfohlene Literatur am Schluss des Berichts) nennt aber folgende wiederkehrende Merkmale von Big Data: Demnach bilden die Grösse und die Komplexität der Datenmenge (Diversität, Dynamik der Zunahme des Datensets etc.) sowie die für deren Analyse verwendeten Technologien die kritischen Faktoren. Gängig ist eine Charakterisierung von Big Data anhand der vier «V»: Volume (Datenmenge), Variety (Heterogenität der Datenquellen und -arten), Velocity (Geschwindigkeit der Zunahme der Datenmenge; Sampling-Rate) und Veracity (Unsicherheit bezüglich der Zuverlässigkeit und des Informationsgehalts der Daten).

Die zunehmende Nutzung von Big Data mag einen qualitativen Sprung hinsichtlich des Umgangs mit Daten bedeuten, doch in rechtlicher Hinsicht beruht deren juristische Bewertung auf Prinzipien, die in den 1970er Jahren entwickelt wurden und die entsprechend kaum für die Herausforderungen von Big Data passen. Dabei ist für Schweizer Unternehmen insbesondere das Datenschutzgesetz (DSG) zu beachten. Werden Daten von in der EU ansässigen Personen bearbeitet, um diesen Dienstleistungen oder Waren anzubieten oder ihr Verhalten zu überwachen (Profiling), kommt zudem die Datenschutz-Grundverordnung der EU (DS-GVO) zur Anwendung. Die DS-GVO gilt auch für Schweizer Unternehmen, wenn sie grenzüberschreitend tätig sind. Die nachfolgende Darstellung beschränkt sich jedoch auf das DSG, wobei dessen Prinzipien weitgehend mit denen des europäischen Rechts deckungsgleich sind.

Die Regeln des Datenschutzrechts kommen zur Anwendung, wenn Personendaten bearbeitet werden. Beide Begriffe – «Bearbeiten» und «Personendaten» – sind im DSG äusserst weit gefasst. Als Bearbeiten wird jeder Umgang mit Personendaten qualifiziert, unabhängig von den angewandten Mitteln oder Verfahren, insbesondere das Beschaffen, Aufbewahren, Verwenden, Umarbeiten, Bekanntgeben, Archivieren oder Vernichten von Daten (Art. 3 lit. e DSG). Als Personendaten gelten alle Angaben, die sich auf eine bestimmte oder bestimmbare Person beziehen (Art. 3 lit. a DSG). Bestimmt oder zumindest bestimmbar ist eine Person, wenn sich ihre Identität aus den Daten selbst, aus dem Kontext oder durch eine Kombination mit weiteren Daten ohne unverhältnismässigen Aufwand feststellen lässt. Als unverhältnismässig gilt der Aufwand, wenn nach der allgemeinen Lebenserfahrung nicht damit zu rechnen ist, dass ein Unternehmen diesen auf sich nehmen wird. Sachdaten und anonymisierte Datensätze, aus denen sich keine Person bestimmen lässt, fallen somit grundsätzlich nicht unter das Datenschutzgesetz. Bei Big-Data-Anwendungen ist in diesem Zusammenhang jedoch stets zu beachten, dass eine Re-Individualisierung mittels Kombination verschiedener Datenbestände mit zunehmender Datenmenge wahrscheinlicher wird und die Daten dann wieder als Personendaten im Sinne des DSG zu qualifizieren sind.

Werden Personendaten bearbeitet, sind die Bearbeitungsgrundsätze des Art. 4 DSG einzuhalten. Gemäss diesen muss bereits die Beschaffung der Personendaten für die betroffene Person erkennbar sein. Ebenso muss der Zweck der Beschaffung ersichtlich sein und die Daten dürfen in der Folge nur für die angegebenen oder aus den Umständen vorhersehbaren Zwecke bearbeitet werden. Die Bearbeitung darf nur rechtmässige Ziele verfolgen, muss sich an den Grundsatz von Treu und Glauben halten und verhältnismässig sein. Verhältnismässig ist die Datenbearbeitung, wenn sie zur Erreichung des angestrebten Zwecks geeignet und erforderlich und den betroffenen Personen zumutbar ist. Letztere Leitlinien manifestieren sich in den Grundsätzen der Datenminimierung und der Speicherbegrenzung. Werden die Bearbeitungsgrundsätze eingehalten, so ist die Bearbeitung zulässig; werden diese Grundsätze nicht eingehalten, ist die Bearbeitung nur dann zulässig, wenn sie sich auf einen Rechtfertigungsgrund nach Art. 13 Abs. 1 DSG stützen kann. Der wichtigste vom Gesetz anerkannte Rechtfertigungsgrund ist die Einwilligung der betroffenen Person. Voraussetzung für eine gültige Einwilligung ist allerdings, dass sie nach angemessener Information und freiwillig erfolgt. Eine Einwilligung kann von der betroffenen Person zudem jederzeit widerrufen werden. In diesem Fall kann die Datenbearbeitung aber immer noch durch ein überwiegendes Interesse des bearbeitenden Unternehmens oder der Öffentlichkeit sowie aufgrund einer gesetzlichen Grundlage gerechtfertigt sein.

Big-Data-Anwendungen stehen in verschiedener Hinsicht im Konflikt mit den Grundsätzen der Erkennbarkeit und der Zweckbindung der Datenbearbeitung. Zunächst verwenden Big-Data-Anwendungen auch automatisiert gesammelte Daten, deren Existenz und Weitergabe (gerade kleinere Unternehmen werden für Big-Data-Analytics häufig spezialisierte Drittunternehmen beiziehen) der betroffenen Person regelmässig nicht bewusst sein kann. Das Beschaffen und Bearbeiten von Personendaten wird in solchen Konstellationen somit gegen den Grundsatz der Erkennbarkeit verstossen. Die Mehrfachverwendung und Neukombination von Daten läuft zudem dem Grundsatz der Zweckbindung diametral entgegen. Das DSG lässt zwar inhaltlich ziemlich weit gefasste Zweckbindungen zu, die den grundlegenden Konflikt aber nicht beheben, sondern nur relativieren können. Die europäische DS-GVO verlangt in diesem Zusammenhang, dass die Einwilligung für einen spezifischen Zweck erteilt werden muss, und ist somit strenger. Bei der Beschaffung müsste die betroffene Person also auf die Kombination von Daten zu Analysezwecken hingewiesen werden, um im Geltungsbereich der DS-GVO gültig einwilligen zu können. Die Zweckbindung ist im Übrigen immer auch von eventuellen Drittbearbeitern einzuhalten, sodass es sich empfiehlt, diesen die Daten nur mit einem Hinweis auf die Zweckbindung zu übergeben.

Da Big-Data-Anwendungen grundsätzlich so viele Daten wie möglich verarbeiten, stehen sie in einem Konflikt mit den Grundsätzen der Datenminimierung und der Speicherbegrenzung. Die Forderung, eine möglichst geringe Menge an Personendaten zu beschaffen, minimiert das Potenzial von Big-Data-Anwendungen, da deren Vorhersagen mit zunehmender Datenmenge (Volume) an Präzision gewinnen. Ebenso vergibt eine konsequente Begrenzung der Speicherdauer durch frühzeitige Anonymisierung oder Löschung der Daten das Potenzial künftiger Nutzungen.

Heute besteht ein Spannungsverhältnis zwischen diesen rechtlichen Vorgaben und dem tatsächlichen Umgang mit Daten in der Informationsgesellschaft. Nicht wenige Juristen halten die geltende Rechtslage für als in weiten Teilen überholt, doch alternative Ansätze fehlen. Vor diesem Hintergrund spielen ethische Überlegungen eine wichtige Rolle, wenn grundlegend über Datenschutz und Privatsphäre in der Informationsgesellschaft nachgedacht werden muss.

Der Abschnitt «Begriffliche und rechtliche Grundlagen» ist mit freundlicher Unterstützung von Prof. Florent Thouvenin und Damian George vom Rechtswissenschaftlichen Institut der Universität Zürich entstanden.

# Big Data in der wissenschaftlichen Literatur

Die bibliometrische Analyse der Fachliteratur und allgemeinen Medien soll einen ersten Eindruck der Bedeutung der Thematik vermitteln. Zu diesem Zweck zeigen wir: 1) die Häufigkeit der publizierten Artikel, die das Stichwort «big data» enthalten; 2) die Verteilung sechs ethischer Themengebiete[1], die in der qualitativen Literaturanalyse identifiziert wurden; 3) die geografische Herkunft hochzitierter Big-Data-Artikel; und 4) die thematische Verteilung der hochzitierten Literatur im Vergleich mit jenen Beiträgen, welche diese Arbeiten zitieren.[2] Das gibt einen Hinweis darauf, dass die Themen «Privatsphäre», «Sicherheit» und «Überwachung» in der wissenschaftlichen Diskussion auf besondere Resonanz stossen. Diese Themen sind nicht vollständig deckungsgleich mit den acht ethischen Gesichtspunkten, anhand derer im weiteren Verlauf des Berichts die Fallbeispiele besprochen werden. Dies erklärt sich dadurch, dass die Themen der Literaturanalyse auf einer explorativen Auswertung der Literatur beruhen. Diese Themen wurden für die ethische Debatte durch Themen ergänzt und verfeinert, die im Rahmen der Expertenworkshops und -interviews identifiziert wurden.

Abbildung 1 zeigt die relative Verteilung der publizierten Artikel im Untersuchungszeitraum in allen drei Datenbanken für die generelle Big Data-Literatur. Generell fällt auf, dass der Ausdruck «big data» vor 2011 praktisch inexistent ist. Der erste Artikel, in dem «big data» im heutigen Verständnis als zusammenhängender Begriff auftaucht, erschien 1998. 2011 beginnt dann ein enorm starkes Wachstum der Zahl von Publikationen. Bemerkenswert ist, dass die Entwicklung in den wissenschaftlichen Datenbanken (WoS und Scopus) jener in den allgemeinen Medien nachhinkt: während sich in Factiva bereits eine Tendenz zur Abflachung zeigt, sind 39,6% (WoS) bzw. 44,7% (Scopus) aller wissenschaftlichen Publikationen zu Big Data im letzten Jahr (2015) erschienen. Diese Tendenz

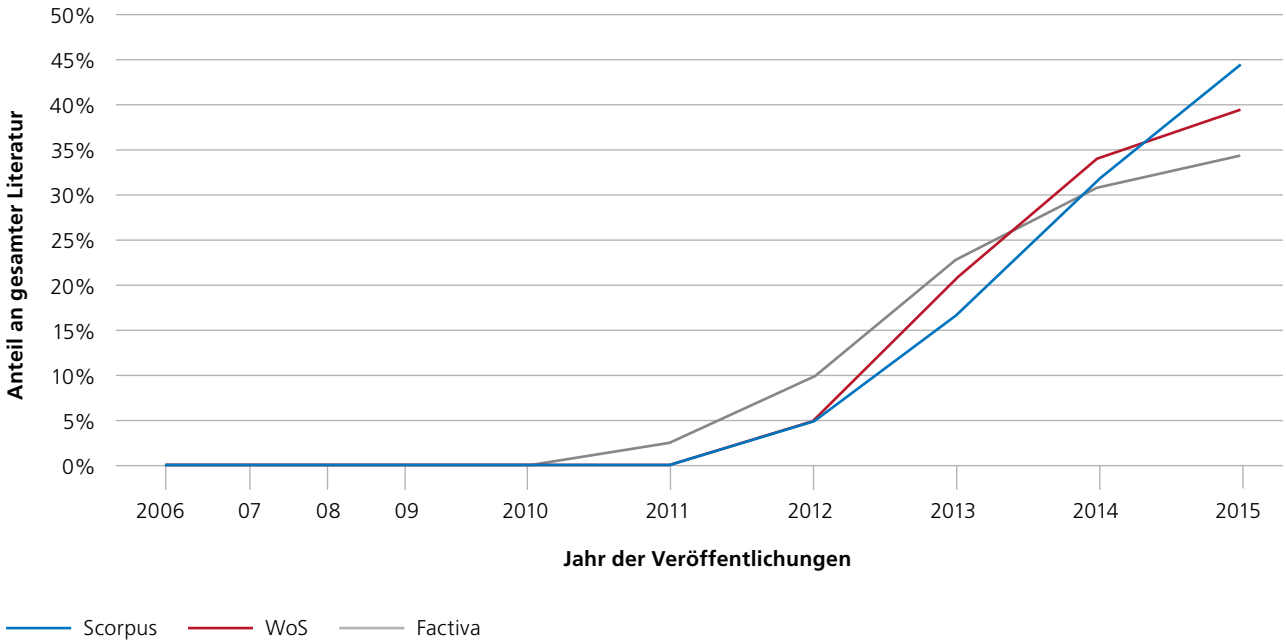



Abbildung 1: Anteil der pro Jahr veröffentlichten Big-Data-Papers gemessen an der Gesamtmenge aller Papers von 2006 bis 2015 pro Datenbank

ist bei den Beiträgen mit ethischer Terminologie noch stärker ausgeprägt – also Beiträgen, welche die in Fussnote 1 (Seite 9) beschriebenen Begriffe enthalten. Bei diesen Big-Data-Ethik-Publikationen zeigt sich zudem, dass der Anteil dieser Literatur pro Jahr an der gesamten Big-Data-Literatur seit 2011 in Factiva stabil bei rund 20% geblieben ist, in den wissenschaftlichen Datenbanken aber von 9,4% auf 13,3% (Wos) bzw. 12,5% auf 26,1% zugenommen hat, d.h. die ethische Debatte hat in der wissenschaftlichen Literatur an Gewicht gewonnen. Die enorme Dynamik ist auch an der Zahl der Zitationen ersichtlich: Im Mai 2015 (dem Zeitpunkt der Voranalyse) wurden die Big-Data-Publikationen von 2006–2015 von rund 3000 anderen Arbeiten zitiert – im März 2016, als die Analysen für die hier dargestellten Abbildungen umgesetzt wurde, generierten die Beiträge desselben Zeitraums mehr als 4000 Zitationen. Die wissenschaftliche Diskussion rund um Big Data trägt somit die typischen Anzeichen eines Hypes.

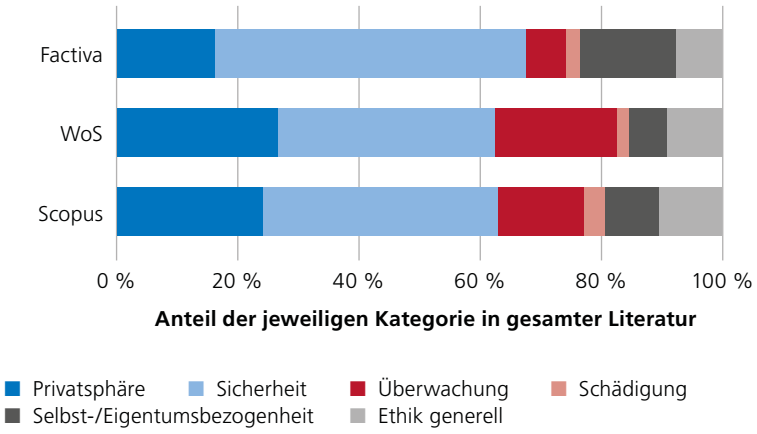



Abbildung 2: Verteilung des Anteils ethischer Themen innerhalb der Literatur. Die Stichworte, welche die ethischen Themen umschreiben, sind in Fussnote 1 (Seite 9) beschrieben.

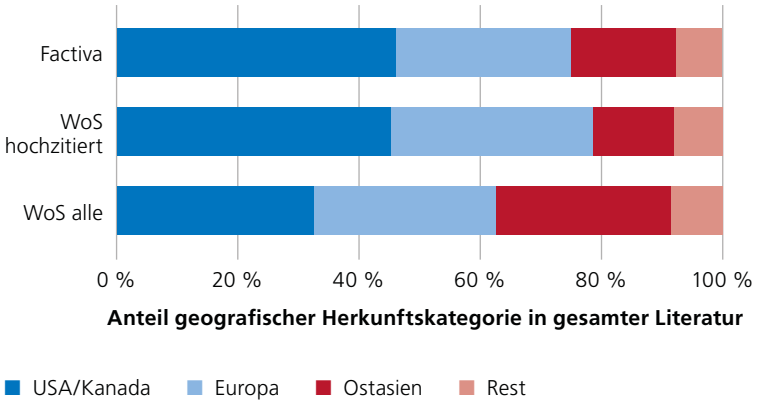



Abbildung 3: Geografische Herkunft der hochzitierten Big-Data-Papers

Abbildung 2 zeigt das relative Gewicht von Publikationen, die jeweils Begrifflichkeiten einer der sechs ausgewählten moralischen Kategorien enthalten. Beim Vergleich der beiden wissenschaftlichen Datenbanken mit den allgemeinen (wirtschaftsfokussierten) Medien zeigt sich insbesondere, dass bei letzteren die Themen «Sicherheit» und «Selbst-/Eigentumsbezogenheit» (also Fragen von Reputation und Eigentum) ein deutlich stärkeres Gewicht einnehmen, während in der wissenschaftlichen Literatur die «klassischen» Themen der Computerethik – Privatsphäre und Überwachung – deutlich dominieren.

Für die weiteren Analysen wurden innerhalb der WoS-Resultate jene Publikationen des Suchzeitraums ausgewählt, welche bis zum Zeitpunkt der Analyse (März 2016) mehr als zehn Zitationen erreicht haben – dies sind insgesamt 164 Beiträge. Abbildung 3 vergleicht den geografischen Ursprung dieser hochzitierten (d.h. wissenschaftlich besonders anerkannten) Beiträge, gemessen an der Heiminstitution der Autoren bzw. im Fall von Factiva der im Artikel vorkommenden Herkunftsbezeichnungen, mit allen des Untersuchungszeitraums sowie mit jenen von Factiva. Dabei zeigt sich, dass der vergleichsweise hohe Anteil (29,0%) von wissenschaftlichen Beiträgen aus dem ostasiatischen Raum bei den hochzitierten Beiträgen deutlich schrumpft (13,5%); hier dominiert die Literatur aus Nordamerika (45,5%).

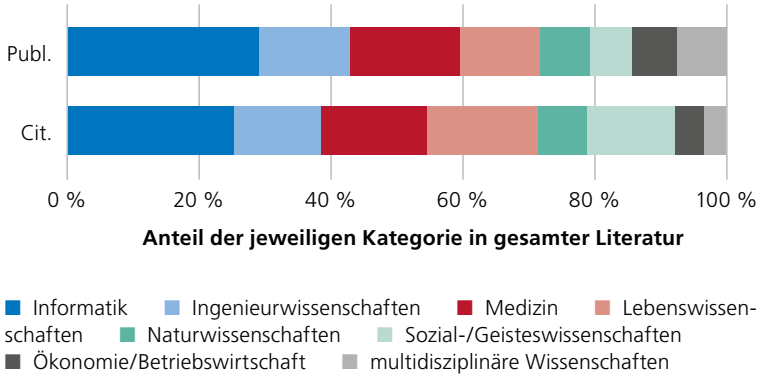



Abbildung 4: Inhaltliche Verteilung der hochzitierten Big Data Paper sowie der diese Arbeiten zitierenden Paper.

Interessant ist schliesslich auch das disziplinäre Profil der hochzitierten Publikationen im Vergleich mit jenen, welche diese Beiträge zitieren. Die Anteile der Lebenswissenschaften (von 12,2% auf 16,5%) und der Sozial- bzw. Geisteswissenschaften (von 6,6% auf 13,2%) nehmen deutlich zu, was ein Hinweis ist, dass in diesen Bereichen die Thematik Big Data besonders Resonanz erhält. Es ist aber nicht auszuschliessen, dass auch Unterschiede in der Zitationskultur der jeweiligen Disziplin zu den genannten Resultaten führen (z.B. dass in den Lebenswissenschaften generell mehr Beiträge zitiert werden). Die Abnahme der multidisziplinären Beiträge (Zeitschriften wie «Nature» und «Science») von 7,3% auf 3,3% entspricht den Erwartungen, während die Abnahme des Bereichs Wirtschaft/Management von 6,9% auf 4,4% darauf hinweist, dass diese Themen quantitativ gesehen weniger stark diskutiert werden. Bei den angegebenen Geldgebern in den hochzitierten Publikationen dominieren öffentliche Geldgeber und Stiftungen aus den USA mit 44,7%, gefolgt von Ostasien mit 28,9% und Europa mit 14,7%. Unternehmen wurden in 8,1% der Fälle als Geldgeber genannt (Rest: 3,5%).

Zusammenfassend ergeben sich aus dieser quantitativen Analyse folgende Schlüsse für diese Arbeit: Erstens ist das Thema Big Data sehr jung und es finden sich Indizien für eine zuweilen medial etwas aufgebauschte Debatte. Zweitens gibt es Anzeichen einer unterschiedlichen Gewichtung ethischer Themen in der wissenschaftlichen und der allgemeinen Literatur. Drittens finden sich Hinweise, wonach Big Data in wissenschaftlicher Hinsicht in den Lebenswissenschaften sowie den Sozial- und Geisteswissenschaften überproportional diskutiert wird. Für die weitere Analyse bedeutet dies, dass (ökonomische) Verheissungen von Big Data mit einer gewissen Vorsicht zu behandeln sind und dass der ethische Fokus über die klassischen Themen Privatsphäre und Überwachung hinaus reichen sollte.

[1] Die Themengebiete wurden durch 2–5 Stichworte wie folgt charakterisiert, wobei die Spezifizität aller Stichwortgruppen vorgängig geprüft wurde: Privatsphäre (privacy OR anonym*), Sicherheit (security OR protection), Überwachung (surveillance OR profiling), Schädigung (discrimination OR harm), Selbst-/Eigentumsbezogenheit (identity OR reputation OR ownership) und Ethik generell (ethic* OR moral OR fairness OR justice OR autonomy). Bei der Suche wurden die Stichwortgruppen dann jeweils mittels des Operators AND mit «big data» verknüpft.

[2] Zu diesem Zweck werden die subject categories des WoS verwendet, die jeweils den Publikationen zugeordnet sind. Diese verweisen auf den disziplinären Hintergrund der Zeitschrift, in der der jeweilige Beitrag veröffentlicht worden ist.

# Big-Data-Anwendungen

**Die Ausführungen in diesem Kapitel illustrieren fünf Anwendungsbeispiele von Big Data, wie sie aktuell von Unternehmen eingesetzt oder künftig möglich sein werden. Diese Einsatzbereiche von Big Data wurden von den Teilnehmenden der Workshops und Interviews identifiziert und diskutiert. Sie dienen als Grundlage für die ethische Diskussion im folgenden Kapitel.**

# Beispiel 1: Debitorenausfälle verhindern

**Um die Zahlungsmoral und das Risiko von Zahlungsausfällen individueller Kunden einzuschätzen, waren Unternehmen bislang auf Daten von Auskunfteien und öffentlichen Verzeichnissen angewiesen. Big-Data-Technologien erlauben Unternehmen, anhand des digitalen Verhaltens ihrer Kunden ihre Kreditwürdigkeitsprüfung zu verfeinern und daraus Rückschlüsse auf die individuelle Bonität zu ziehen.**

Für die Bonitätsbewertung ihrer Kunden greifen Unternehmen traditionell auf die Dienste klassischer Auskunfteien wie der schweizerischen Zentralstelle für Kreditinformation (ZEK) oder der deutschen Schutzgemeinschaft für allgemeine Kreditsicherung (SCHUFA) zurück, welche z.B. auf Basis von Konsumkredit-, Leasing- und Kreditkartenverpflichtungen Bonitätsinformationen über private Haushalte erstellen. Sie beziehen zudem Informationen aus öffentlichen Verzeichnissen (z.B. über Insolvenzverfahren oder Betreibungen) sowie soziodemografische Daten mit ein, um einen aussagekräftigen Scoring-Wert zu ermitteln, der angibt, mit welcher Wahrscheinlichkeit Rechnungen bezahlt werden.

Social-Scoring-Verfahren, die auf Big Data beruhen, bieten Unternehmen neue Möglichkeiten, die Zahlungsmoral eines individuellen Kunden einzuschätzen. Hierzu analysieren Algorithmen z.B. Einträge und Verhalten des Kunden auf sozialen Medien: beruflicher Werdegang, Arbeitgeber, Freizeitbeschäftigungen, Ferienziele, Freunde und Likes. Darüber hinaus fliessen Informationen in die Scoring-Modelle ein, die auf den ersten Blick wenig mit dem Risiko von Zahlungsausfällen zu tun haben, z.B. solche bezüglich Such- und Surfverhalten, verwendete Schriftarten, die Art und Weise, wie das jeweilige Formular ausgefüllt wird (Tempo und Rhythmus, Häufigkeit der Betätigung der Lösch- oder Copy-Paste-Taste, Rechtschreibfehler etc.), sowie technische Daten des verwendeten Computers (z.B. Alter und Preis).

Im E-Commerce findet Social Scoring bereits Anwendung. Der Kauf auf Rechnung ist eine der beliebtesten Zahlungsarten bei Online-Einkäufen. Bei dieser Zahlungsart sind die Händler einem erhöhten Risiko von verspäteten Zahlungen und Zahlungsausfällen ausgesetzt. Daher setzen sie im Rahmen des Bestellvorgangs eine aktive Zahlungsartensteuerung ein, die auf Social-Scoring-Verfahren beruht. Während des Bestellungsvorgangs leitet der Algorithmus in Echtzeit aus den verfügbaren Datenpunkten einen Scoring-Wert ab. Wird einem Kunden ein positiver Scoring-Wert zugeschrieben, erhält er die Auswahl zwischen diversen Bezahlmethoden, inklusive Kauf auf Rechnung. Fällt das Scoring negativ aus, stehen dem Kunden nur Zahlungsarten wie Vorauskasse, Lastschrift oder Kreditkarte zur Verfügung, die für den Händler ein geringeres Risiko aufweisen.

Auch im Bankwesen gewinnt Social Scoring zunehmend an Bedeutung. Verschiedene digitale Banken bieten Kredite unter der Voraussetzung an, dass der Antragsteller weitgehenden Zugang zu persönlichen Daten gewährt. Die technische Voraussetzung hierfür ist, dass der Kunde eine App herunterlädt, die z.B. Informationen zu Aufenthaltsorten, Telefonanrufen (Anzahl, Dauer und Uhrzeit), Anzahl erhaltener und versandter SMS, Adressbüchern (z.B. ob die Kontakte ohne Nachnamen gespeichert sind) oder Akkulaufzeit an den Finanzdienstleister überträgt. Je mehr Daten ein Antragsteller freigibt, desto höher die Chancen auf einen Kredit, denn desto treffsicherer kann die Bank seine Kreditwürdigkeit bewerten und das Risiko eines Debitorenausfalls minimieren. Aufgrund der technischen Entwicklung ist zu erwarten, dass der Zugang zu Finanzdienstleistungen zukünftig noch stärker als heute davon abhängt, welche Daten der einzelne Kunde der Bank freigibt.

# Beispiel 2: Risikomanagement verbessern

**Mithilfe ihres Risikomanagements versuchen Unternehmen, frühzeitig Risiken für ihre Organisation zu erkennen, zu bewerten und zu kontrollieren. Insbesondere für die Versicherungswirtschaft, für die die Risikoabschätzung den Kern der Geschäftstätigkeit darstellt, bieten Big-Data-Anwendungen ein grosses Potenzial, Risiken besser zu erkennen und schneller, günstiger und genauer zu bewerten.**

Ein zentrales Anwendungsfeld von Big Data im Risikomanagement von Versicherern liegt in der Festsetzung der Prämien. Um deren Höhe zu bestimmen, prognostizieren Versicherer traditionell die Wahrscheinlichkeit und das Ausmass eines Schadens für eine bestimmte, möglichst homogene Gruppe von Individuen während einer bestimmten Versicherungsperiode. Beispielsweise berechneten Versicherer auf Basis von Erfahrungswerten und Versicherungsforderungen, dass junge Männer häufiger in schwere Autounfälle verwickelt sind als andere demografische Gruppen. Dies führte dazu, dass männliche Junglenker höhere Prämien für Fahrzeugversicherungen bezahlen.

Big Data ermöglicht Versicherern nun, neue und komplexere Datenbestände zu analysieren, die Eintrittswahrscheinlichkeit eines Versicherungsfalls mit höherer Treffsicherheit vorherzusagen und so die Risikoklassifizierung zu verfeinern. Gespeist werden diese Datenbestände u.a. von immer zahlreicheren Sensoren, die in Maschinen verbaut oder von Menschen getragen werden.

Einer der bekanntesten Anwendungsbereiche von Big Data in der Versicherungswirtschaft stellt der Einsatz von Telematik zur Erfassung und Analyse von Echtzeitdaten in Autos dar. Informationen zum Fahrstil (z.B. Beschleunigungs- und Bremsverhalten) sowie zu den Fahrten (z.B. Dauer, Länge und Art der Strecke) werden an die Versicherungsgesellschaft übermittelt, die darauf basierend die Wahrscheinlichkeit eines Unfalls oder Diebstahls berechnet und die Versicherungsprämie entsprechend anpasst («pay as you drive»).

Auch für Kranken- und Lebensversicherungen liefern immer mehr Sensoren wie Smart Watches oder Fitnessarmbänder Daten für eine personalisierte Risikoklassifizierung der Versicherungsnehmer. Diese Daten ermöglichen Rückschlüsse auf Gewohnheiten, Lebensstil, Aktivitätsgrad und Gesundheit des Trägers. Versicherer im Ausland bieten bereits Versicherungsmodelle an, die an solche Aktivitätstracker geknüpft sind. Die Kunden können anhand der damit generierten Daten einen gesunden Lebenswandel belegen oder aufzeigen, dass sie ungesunde Verhaltensmuster verbessern, um eine Prämienreduktion zu erreichen. Dereinst können auch implantierte Mikrochips solche Informationen liefern. Bereits heute nutzen weltweit mehrere tausend Menschen RFID-Implantate, um Haus- und Autotüren zu öffnen, Kopierer zu bedienen, kontaktlos zu bezahlen, Passwörter und Notfalldaten zu speichern und Kontaktdaten auszutauschen.

Eine weitere Datenquelle stellen die menschlichen Erbgut-Informationen dar. Mithilfe von Genomsequenzierung kann die genetisch bedingte Wahrscheinlichkeit für das Auftreten von Krankheiten ermittelt werden. Das eröffnet technisch die Möglichkeit, dass Versicherungsgesellschaften je nach Risikoprofil die Prämien individuell festsetzen oder für bestimmte erwartete Krankheitsbilder den Versicherungsschutz ausschliessen können, auch wenn diese zum Zeitpunkt des Versicherungsabschlusses nur aufgrund der genetischen Disposition prognostiziert werden.

Ein Anwendungsbereich, in dem Versicherer bereits heute stark auf Big-Data-Anwendungen setzen, ist das Schadenmanagement (Claims Management), und hierbei insbesondere die Aufdeckung von Versicherungsbetrug. Mithilfe von Profiling- und Predictive-Modelling-Verfahren werden Variablen aus Schadensmeldungen mit einer Betrugsfall-Datenbank abgeglichen. Dieser Abgleich berücksichtigt einerseits das Verhalten des Anspruchstellers gegenüber der Versicherung, andererseits werden zunehmend Daten involvierter Partneragenturen wie Reparaturwerkstätten sowie weiterer öffentlich zugänglicher Quellen, z.B. soziale Medien, hinzugezogen.

## Beispiel 3: Angebotskonditionen massschneidern

**Traditionell berechnen Unternehmen den für weite Kundengruppen identischen Produktpreis anhand ihrer Kosten, der Preise der Mitbewerber oder dem aggregierten Kundenverhalten. Lange galten auf einzelne Käufer massgeschneiderte Preise als nicht realisierbar, da das Wissen über die Zahlungsbereitschaft einzelner Kunden fehlte – Big Data ermöglicht es Unternehmen prinzipiell, den idealen Preispunkt für jeden Kunden individuell zu bestimmen.**

Dynamische Preise, die auf aggregierten Echtzeitdaten über Faktoren auf Angebots- und Nachfrageseite basieren, gehören in bestimmten Branchen seit Längerem zum Alltag. Nicht nur Fluggesellschaften, sondern auch Online-Händler oder Tankstellenbetreiber setzen Algorithmen ein, die in Abhängigkeit zahlreicher Faktoren wie Nachfrage, Verfügbarkeit, Wetter, Uhrzeit oder Konkurrenzverhalten die Preise für ihre Produkte dynamisch bestimmen und fortlaufend anpassen.

Der nächste Schritt sind individualisierte Preise. Diese rücken näher an den einzelnen Käufer heran und ermöglichen eine noch zielgenauere Abschöpfung der Konsumentenrente. Zur Preisbestimmung werden hierbei nicht nur Umweltfaktoren einbezogen, sondern die Kunden als Individuen erfasst. Daten wie beispielsweise Geschlecht, Alter, Wohnort, Einkommen, Verwandtschafts- und Freundschaftsbeziehungen, Beschäftigungsstatus und Tätigkeitsgebiet, Bewegungsmuster sowie persönliche Vorlieben und Wertvorstellungen können in die Preisbildung einfliessen. Möglich machen dies Cookies, Kundenkarten, Smartphone-ID, SIM-Kartennummer, Bluetooth-Kennung, Adresse des WLAN-Moduls, GPS oder IP-Adressen. Dank dieser Technologien können persönliche Merkmale und Verhaltensweisen erfasst und mit zunehmender Treffsicherheit einem Individuum zugeordnet werden.

Dazu zwei Illustrationen: Ein Online-Reiseportal sorgte unlängst für Aufsehen, da Apple-Nutzern teurere Hotelzimmer angezeigt wurden als Windows-Nutzern. Das Portal hatte herausgefunden, dass Apple-Nutzer im Schnitt 30% mehr für ihren Aufenthalt ausgeben. Ein anderes Beispiel: 2015 erhielten Kunden eines Onlineshops personalisierte Bons, die Rabatte auf Produkte gewährten, die das Unternehmen auf der Basis des Einkaufsverhaltens des individuellen Kunden ausgewählt hatte. Die Kombination gekaufter Produkte und soziodemografischer Daten liefert Hinweise auf Produktvorlieben und Zahlungsbereitschaft. Bestellt jemand regelmässig Bier der günstigen Eigenmarke, könnte er mit Rabatt-Gutscheinen dazu bewegt werden, längerfristig auf ein teureres Markenbier umzusteigen.

Diese Beispiele zeigen: Je mehr Daten die Kunden hinterlassen und je zuverlässiger diese dem einzelnen Individuum zugeordnet werden können, desto präziser können Algorithmen die individuelle Zahlungsbereitschaft bestimmen und die Unternehmen ihre Preise entsprechend anpassen. Während individualisierte Preise im Online-Handel bereits teilweise realisiert sind, steht die Entwicklung in der nicht-digitalen Welt noch in den Kinderschuhen. Technologisch ist aber denkbar, dass Kunden in Zukunft mithilfe von Kameras, iBeacons und Gesichtserkennungssoftware auch im stationären Handel identifiziert und entsprechend ihrer Zahlungsbereitschaft individuell angesprochen werden könnten.

## Beispiel 4: Effizienz der Werbemassnahmen steigern

**Werbung ist dann erfolgreich, wenn sie die richtigen Personen erreicht. Während traditionelle Werbung auf eine breite Streuung ihrer Botschaft angewiesen ist, erreicht auf Big Data gestützte Werbung mit höherer Treffsicherheit jene Verbraucher, zu deren Interessen sie wirklich passt. Big Data ermöglicht Unternehmen somit, durch die gezielte Ansprache potenzieller Kunden die Effizienz ihrer Werbemassnahmen zu steigern.**

Traditionell setzen werbetreibende Unternehmen in erster Linie auf aufwändige und teure Kampagnen, zum Beispiel in TV, Zeitschriften, Zeitungen oder Radio. Nach dem Giesskannenprinzip wird versucht, möglichst viele Personen zu erreichen, wobei nur wenige Informationen zu deren Demografie vorliegen. So erreicht beispielsweise klassische Plakatwerbung an Bahnhöfen die Zielgruppe der Pendler, die aber nicht nach weiteren Merkmalen unterschieden werden können.

Big Data bildet die Grundlage für einen Paradigmenwechsel in der Werbeindustrie. Durch die Analyse von Daten aus Suchverläufen, sozialen Netzwerken, GPS-Daten und vielen weiteren digitalen Anwendungen können Unternehmen ermitteln, wie, wo, wann und was Kunden einkaufen, wie viel sie

ausgeben und was sie in ihrer Freizeit tun. Targeting lautet das Zauberwort: Dieses ermöglicht, die richtigen Personen zum richtigen Zeitpunkt am richtigen Ort anzusprechen.

Internetnutzern längst bekannt ist das Prinzip des Re-Targetings: Ein Kunde, der sich im Onlineshop Schuhe anschaut, wird nach Verlassen der Seite von diesen Schuhen regelrecht verfolgt – unabhängig davon, ob er diese gekauft hat oder nicht. Das von vielen Kunden als belästigend empfundene Re-Targeting wird zunehmend durch Pre-Targeting abgelöst. Mithilfe von «Predictive Analytics» werden anhand des individuellen Such- und Kaufverhaltens Vorhersagen über weitere für den Kunden relevante, ergänzende Angebote getroffen, etwa die passende Pflegecrème zum schwarzen Glattlederschuh. Werber ziehen zudem Geo-Targeting heran, welches ihnen anhand der IP-Adresse ermöglicht, nur die Nutzer einer gewünschten Region anzusprechen. Für ein Location-Based Marketing werden mittels Sendemastinformation oder GPS-Koordinaten noch genauere Informationen über den Aufenthaltsort einer Person und ihre Bewegungsmuster ausgewertet.

Einen Schritt weiter geht das Emotional Targeting. Selbstlernende Algorithmen erfassen mithilfe von Kameras und Mikrofonen Gesichtsausdrücke und Stimmen und versuchen in Echtzeit die Gefühlslage der beobachteten Personen zu analysieren. Für diese Anwendung wurden Millionen Gesichtsausdrücke und Stimmlagen in zahlreichen Ländern kategorisiert. Die Technologie wurde bereits bei onlinebasierten Computerspielen mit dem Ziel erprobt, den Spielern im geeigneten Moment personalisierte Werbung einzublenden. Gefühle spielen in der wirtschaftlichen Entscheidungsfindung eine zentrale Rolle. Emotional involvierte Konsumenten beachten, erinnern und kaufen das Produkt mit höherer Wahrscheinlichkeit.

Ferner gibt es Bestrebungen, personalisierte Werbung in der nicht-digitalen Welt einzusetzen. iBeacons, WLAN- und GPS-Daten ermöglichen es, Individuen präzise zu orten und ihnen lokal relevante, personalisierte Werbung zuzuspielen. So ist es technisch vorstellbar, Aussenwerbestationen auf die persönlichen Präferenzen der Passanten zu konfigurieren: Passiert ein einsamer Single ein digitales Werbeplakat, wird auf diesem Werbung für einen nahegelegenen Singletreff eingeblendet.

## Beispiel 5: Neue Umsatzquellen erschliessen und Innovationen hervorbringen

**Big-Data-Anwendungen bieten Unternehmen das Potenzial, mithilfe ihrer Daten neue Umsatzquellen zu erschliessen, ihr bestehendes Produkt- und Dienstleistungsportfolio zu verbessern und Innovationen hervorzubringen.**

Laut einer aktuellen Studie sehen viele der befragten Unternehmen in Big-Data-Anwendungen eine Umsatzquelle, die für das Unternehmen genauso wertvoll werden wird wie ihr traditionelles Kerngeschäft. Insbesondere der Datenhandel bietet grosses kommerzielles Potenzial.

Hierzu drei Beispiele: Mobile Endgeräte und die darauf installierten Apps ermöglichen Einblicke in die Persönlichkeit und den Alltag der Nutzer. Weit über die Hälfte der Apps übertragen Daten wie Alter, Geschlecht, Standort, Adressbuch, Kalender und Geräte-ID an Drittfirmen. Ein weiteres Beispiel: Mobilfunkbetreiber verfügen über detaillierte geografische (z.B. Aufenthaltsort) und soziodemografische (z.B. Geschlecht und Alter) Daten ihrer Kunden, die sie vermehrt anderen Unternehmen und staatlichen Institutionen zur Verfügung stellen. Auch Sportartikelhersteller finden in der Quantified-Self-Bewegung – dem Selbstvermessungstrend – eine neue Umsatzquelle. Sie werden zunehmend zu Herstellern von Wearable Technologies, indem sie in ihre Produkte, wie z.B. Sportschuhe, T-Shirts oder Armbänder, Sensoren integrieren. Diese Entwicklung führt dazu, dass Unternehmen zwecks Zugang zu Daten zwar ihr ursprüngliches Angebot beibehalten, z.B. das Erbringen von Telekommunikationsdienstleistungen oder den Verkauf von Sportartikeln. Zunehmend stellen aber nicht mehr diese Produkte und Dienstleistungen, sondern die damit gewon-

nenen Daten und deren Kommerzialisierung das (finanzielle) Herzstück des Geschäftsmodells dar.

Immer wichtiger werden Big-Data-Anwendungen auch für die Angebotsoptimierung und im Innovationsprozess. Hierzu einige Beispiele: In Autos installierte Telematiksysteme liefern Automobilherstellern Daten über verschiedene Parameter wie Fahrverhalten, Verschleiss oder Verbrauch. Basierend auf diesen Daten können die Hersteller mithilfe von Predictive Analytics beispielsweise Schwachstellen neuer Modelle erkennen und beheben sowie After-Sales-Angebote (z.B. Wartung und Reparatur) optimieren. Daten aus sozialen Netzwerken, Autoblogs und Videoaufzeichnungen aus Showrooms und von Messen unterstützen sie dabei, Trends, Kundenbedürfnisse, Nutzungsgewohnheiten und Vorlieben zu identifizieren. Diese Erkenntnisse fliessen in die Fahrzeugentwicklung, die Produktionsplanung und das Qualitätsmanagement ein.

Auch Streaming-Dienste setzen für die Produktion eigener Filme und Serien ihr datengestütztes Wissen über Zuschauerinteressen und Konsumverhalten ein. Auf der Basis dieser Erkenntnisse treffen sie Entscheidungen bezüglich Genre, Handlung und Schauspieler und versuchen so den kommerziellen Erfolg ihrer Eigenproduktionen sicherzustellen.

Ein weiteres Beispiel: Spracherkennungssoftware, die in der Automobil- und Unterhaltungsindustrie sowie in der medizinischen Verwaltung genutzt wird, nutzt Big-Data-Anwendungen, um gesprochene Sprache immer besser verarbeiten zu können. Hierzu wird das Gesprochene an einen Server übertragen. Zusätzlich werden orts- und zeitspezifische Daten gespeichert. Daten wie Arbeits- und Wohnort, Alter, Geschlecht oder Bildungshintergrund helfen der Software, den Kontext des Sprechers einzubeziehen und damit die Spracherkennungsqualität zu verbessern.

Überdies eröffnen Big-Data-Anwendungen neue Möglichkeiten in der Infrastrukturplanung und -optimierung sowie im Verkehrsmanagement. So können Mobilfunkdaten z.B. genutzt werden, um Informationen über Kundenströme durch Einkaufszentren oder Strassenverkehrsströme zu erfassen (z.B. Verkehrsdichte, Fahrgeschwindigkeit, Reisezeit pro Streckenabschnitt). Diese Erkenntnisse dienen beispielsweise dazu, ein vertieftes Verständnis über Stauentwicklung zu erlangen und die Verkehrslenkung zu verbessern.

# Ethische Debatte

Ethische Bewertungen wirtschaftlicher Aktivitäten stehen in der Gefahr, einen Gegensatz zwischen Ethik und Wirtschaft zu postulieren. Ethik bestimmt gemäss diesem Verständnis die Grenzen des wirtschaftlich Erlaubten. Das ist sicher eine wichtige Funktion von Ethik; allerdings dürfen drei Aspekte nicht unterschlagen werden:

1. Vielen wirtschaftlichen Argumentationen liegt die Annahme zugrunde, dass möglichst wenig regulierte Märkte insgesamt moralisch wertvolle Ziele (z.B. ökonomische und soziale Wohlfahrt) erreichen helfen. Unbestritten ist aber auch, dass Märkte bestimmter politisch gesetzter Rahmenbedingungen bedürfen, um diese Ziele wirklich zu erreichen. Wie weit diese ethisch begründete staatliche Regulierung gehen soll, ist allerdings umstritten.

2. Abstrakte ethische Normen und Werte sind zwar meist unbestritten, aber was sie bezüglich konkreter Aktivitäten in der Praxis bedeuten, ist oft nicht eindeutig.

3. Zudem müssen die mit den Normen und Werten verbundenen Ansprüche (z.B. diejenigen der Konsumenten) oft gegen andere, ebenso berechtigte Ansprüche (z.B. diejenigen der Unternehmen) abgewogen werden.

In den durchgeführten Workshops identifizierten und diskutierten die Experten acht ethische Normen und Werte, die von Big-Data-Anwendungen berührt werden. Im Folgenden werden diese kurz vorgestellt, sofern möglich auf die oben beschriebenen fünf Big-Data-Anwendungen bezogen und kritisch diskutiert.

## Schutz der Privatsphäre

Gemäss Art. 12 der allgemeinen Erklärung der Menschenrechte darf niemand «willkürlichen Eingriffen in sein Privatleben, seine Familie, sein Heim oder seinen Briefwechsel noch Angriffen auf seine Ehre und seinen Beruf ausgesetzt werden. Jeder Mensch hat Anspruch auf rechtlichen Schutz gegen derartige Eingriffe oder Anschläge.» Ziel dieser Bestimmung ist es, Lebensbereiche der Individuen zu schützen, in denen diese sich frei bewegen, entwickeln und verhalten können.

In der bisherigen Datenschutzdebatte wird dieser Wert mit den beiden Prinzipien «Zweckbindung» und «Datensparsamkeit» geschützt: Daten sollen nur zu dem Zweck eingesetzt werden, für den sie erhoben wurden, und es sollen nur so viele Daten erhoben werden, wie für einen bestimmten Zweck tatsächlich benötigt werden. Es leuchtet ein, dass eine Routenplaner-App Zugriff auf die Standortdaten des Nutzers benötigt, um ihren Zweck zu erfüllen. Viele Apps verletzen jedoch das Prinzip der Sparsamkeit, indem sie von den Nutzern Zugriff auf Daten verlangen, die für die Erfüllung der eigentlichen Anwendung nicht erforderlich sind. Beispielsweise, wenn die Nutzung einer Taschenlampen-App den Zugriff auf die Standortdaten voraussetzt.

Der klassische Datenschutz und seine beiden Prinzipien werden offensichtlich in allen fünf genannten Big-Data-Anwendungen regelmässig in Frage gestellt. Im Kontext zielgerichteter und personalisierter Werbung (Beispiel 4) mögen einzelne Merkmale nicht sehr privat erscheinen, doch aus der Kombination verschiedener Merkmale können sehr private Rückschlüsse gezogen werden. Eine Kundin stört es kaum, dass der Detailhändler weiss, welche Bodylotion sie kauft; wenn aber der Detailhändler aus der Kombination der gekauften Produkte mit hoher statistischer Wahrscheinlichkeit auf eine Schwangerschaft schliessen kann, greift dies empfindlich in die Privatsphäre ein.

Weitere Risiken für die Privatsphäre ergeben sich, wenn über einzelne Unternehmen hinaus verschiedene Datenbestände verknüpft und neue Erkenntnisse daraus gewonnen werden. Doch gerade in der Kombination verschiedenartiger Daten und in ihrer Verwendung zu anderen Zwecken als denjenigen, zu denen sie ursprünglich erhoben wurden (Sekundärnutzung), liegt das grosse Potenzial von Big Data. Die Mehrfachverwendung und Neukombination von Daten steht im Gegensatz zur Zweckbindung. Ausserdem stösst der Datenschutz bei der Integration verschiedener Datenbestände an seine Grenzen. Eine Anonymisierung der Daten ist für den Schutz der Privatsphäre nicht ausreichend, denn aus der Kombination verschiedener anonymisierter Datenbestände sind oftmals Rückschlüsse auf Personen möglich.

Im Falle des Verhinderns von Debitorenausfällen (Beispiel 1) und des Verbesserns des Risikomanagements (Beispiel 2) müssen die erwähnten Eingriffe in die Privatsphäre allerdings mit den berechtigten Ansprüchen von Unternehmen abgewogen werden. Niemand bestreitet, dass Unternehmen ein Recht darauf haben, über die Bonität ihrer Debitoren und andere geschäftsrelevante Risiken Bescheid zu wissen. So gesehen ist nicht jeder Eingriff in die Privatsphäre ethisch gleichermassen problematisch. Welche Eingriffe in die Privatsphäre ethisch erlaubt sind und welche nicht, muss im Einzelfall abgewogen werden.

Wenn im Hinblick auf das Massschneidern von Angebotskonditionen (Beispiel 3) Angaben aus sehr unterschiedlichen Bereichen (z.B. Angaben über Finanztransaktionen, Bonität, medizinische Behandlungen, privaten Konsum, soziale Beziehungen, Berufstätigkeit) zusammengeführt werden, stellt dies in seiner Gesamtheit einen Eingriff in die Privatsphäre dar. Darüber hinaus erweisen sich bei diesen Anwendungen auch die fehlende Transparenz, die allfällige Diskriminierung und die Verletzung der kontextuellen Integrität der Daten als ethische Herausforderungen. Diese Aspekte werden im weiteren Verlauf des Berichts diskutiert.

Dass Unternehmen mit Big Data die Effizienz ihrer Werbemassnahmen steigern (Beispiel 4) sowie neue Umsatzquellen erschliessen und Innovationen hervorbringen wollen (Beispiel 5), sind grundsätzlich legitime wirtschaftliche Anliegen. In diesen Fällen müssen die für die Kunden generierten Mehrwerte (z.B. gezieltere Werbung) mit den Nachteilen (z.B. kommerzielle Überwachung durch Unternehmen) abgewogen werden. Aus ethischer Perspektive können Eingriffe in die Privatsphäre unter der Voraussetzung gerechtfertigt sein, dass die Kunden transparent über das Ausmass der Datenerhebung informiert sind, zu diesen Datenerhebungen ihre Zustimmung gegeben haben und – falls sie nicht einverstanden sind – realistische Alternativen erhalten. Das Problem der Verständlichkeit der so genannten Allgemeinen Geschäftsbedingungen (AGB) wird im Abschnitt «Transparenz» diskutiert.

## Gleichheit und Nichtdiskriminierung

Unter Diskriminierung wird in der Regel eine Ungleichbehandlung von Personen verstanden, die sachlich nicht gerechtfertigt ist. Nichtdiskriminierung gehört zu den Fairnessgeboten, die in unserer Gesellschaft unbestritten sind und sich z.B. im Grundsatz der Gleichheit vor dem Gesetz widerspiegeln. Dass Personen ungleich behandelt werden (z.B. aufgrund unterschiedlicher Leistungen unterschiedliche Löhne erhalten) ist oft unvermeidbar bzw. gilt als allgemein akzeptiert. Ethisch problematisch wird es aber, wenn dabei Kriterien eine Rolle spielen (z.B. die Hautfarbe, das Geschlecht oder die Religion), welche im Hinblick auf den Zugang zu bestimmten Gütern, Chancen und Positionen nicht relevant sind.

Dies kann bei individualisierten Preisen (Beispiel 3) geschehen: Algorithmen ordnen Individuen, basierend auf z.T. nicht oder nur beschränkt beeinflussbaren Merkmalen verschiedenen Klassen zu, die als Basis für den individualisierten Preis dienen. Der ethischen Brisanz individualisierter Preise scheinen sich auch Unternehmen bewusst zu sein. Sie experimentieren bislang vor allem mit individualisierten Rabatt-Gutscheinen und nicht mit individualisierten Preisen, was aber im Ergebnis kein entscheidender Unterschied ist. Big-Data-Anwendungen bieten den Unternehmen die Aussicht, die Konsumentenrente mittels individualisierter Preise noch besser abzuschöpfen und ihre Umsätze zu maximieren. Kunden werden dabei bezüglich Preis nicht gleich, sondern systematisch unterschiedlich behandelt.

Dagegen, dass in die Preisgestaltung die Zahlungsbereitschaft der Kunden einfliesst, ist nichts einzuwenden, solange unverschuldete Notlagen nicht ausgenutzt werden und keine Monopole bestehen. Institutionen wie der Basar oder die Auktion, wo Verkäufer und Käufer um einen Preis feilschen bzw. mitbieten, um die Zahlungsbereitschaft zu eruieren, sind weithin akzeptierte Formen wirtschaftlicher Interaktion. Aufgrund von Big Data kann es jedoch zu Informationsasymmetrien kommen, die die bezüglich klassischer Märkte postulierte Eigenschaft effizienter Allokation

aushebeln können. Ethisch relevant erscheint in diesem Zusammenhang ein Aspekt zu sein: Die Kunden wissen nicht, inwiefern und aufgrund welcher Kriterien sie möglicherweise diskriminiert werden (vgl. Abschnitt «Transparenz»).

Individualisierte Preise erscheinen besonders dann problematisch, wenn sie innerhalb einer Branche flächendeckend eingesetzt werden und Konsumenten nicht mehr auf andere Anbieter ausweichen können. Allerdings wäre dieses Argument nur dann entscheidend, wenn man nicht mehr zwischen unterschiedlichen individualisierten Preisen wählen könnte. Eine solche Monopolsituation bestünde erst dann, wenn jeder Anbieter denselben Algorithmus sowie dieselbe Informationsbasis für den jeweiligen Kunden zur Bestimmung des Preises verwenden würde. Dann hätte man es mit einer (neuen) Form eines Monopols oder eines Kartells zu tun, also mit dem klassischen Problem eines Marktes ohne Wettbewerb.

Ein weiteres Beispiel: Aufgrund von Social Scoring erhalten auch Menschen aus weniger privilegierten Schichten Zugang zu Krediten, denen der Zugang zum klassischen Bankgeschäft bislang verwehrt blieb. Um Debitorenausfälle zu verhindern (Beispiel 1), verlangen digitale Finanzdienstleister sehr persönliche Informationen. Ansonsten erhalten die Antragsteller keinen Zugang zu Finanzdienstleistungen. Im klassischen Bankgeschäft werden derart private Einblicke nicht verlangt. Handelt es sich dabei um einen Verstoss gegen Regeln der Fairness, also um eine unzulässige Diskriminierung? Wie im Abschnitt «Schutz der Privatsphäre» ausgeführt, ist die Ungleichbehandlung von Kreditnehmern im Hinblick auf das Vermeiden von Debitorenausfällen ein berechtigtes Anliegen von Kreditgebern. Dieses Anliegen muss aber gegen das ebenso berechtigte Interesse der Kreditnehmer, ihre Privatsphäre zu schützen, abgewogen werden. Problematisch wäre zudem, wenn für das Risikomanagement Daten eingesetzt werden, die in anderen Kontexten oder zu anderen Zwecken erhoben wurden, also die kontextuelle Integrität der Daten oder das Prinzip der Zweckbindung verletzt würde.

## Informationelle Selbstbestimmung

Informationelle Selbstbestimmung wird in der Regel definiert als das Recht des Einzelnen, selbst über das Erheben, Speichern, Verwenden und Weitergeben persönlicher Daten bestimmen zu können. Sie wurzelt im Grundsatz der Autonomie, die Menschen Rechte (und Pflichten) hinsichtlich der Gestaltung und Kontrolle des eigenen Lebens gibt. Ein praktischer Ausdruck von Autonomie und Selbstbestimmung besteht im Prinzip der informierten Zustimmung, das ursprünglich aus dem medizinischen Bereich stammt. Im Fall von Daten ist damit gemeint, dass eine Person der Verwendung ihrer persönlichen Daten explizit und im Wissen, was mit diesen Daten geschehen wird, zustimmen sollte.

Ein Problem wäre besonders dann gegeben, wenn im Rahmen zielgerichteter und dadurch effizienterer Werbung (Beispiel 4) Daten über Personen aus unterschiedlichen Quellen und Kontexten so aufbereitet werden, dass Personen gezielt emotional manipuliert werden können. Die Werbung also nicht nur generelle emotionale Trigger

verwendet (z.B. Stereotype von Freiheit, Attraktivität oder Individualität in der Tabakwerbung), sondern gezielt den emotionalen Zustand einer Person erfasst und darauf abgestimmte Botschaften vermittelt. Problematisch ist daran einerseits, dass die Betroffenen in der Regel ihre Zustimmung zu diesem Verwendungszweck der Daten nicht bewusst gegeben haben (vgl. Abschnitt «Transparenz»), und dass andererseits aufgrund des manipulativen Charakters die Möglichkeiten einer freien Meinungsbildung und einer informierten Zustimmung zu einem Produkt oder einer Dienstleistung eingeschränkt sind, was letztlich die Wahlfreiheit der Personen beeinträchtigt. Dass Werbung manipulative Züge aufweisen kann, ist kein neues Phänomen, und bewusste Täuschungen sind in keinem Fall gerechtfertigt. Die Möglichkeiten ethisch problematischer Manipulation durch Werbung (z.B. durch das sogenannte Emotional Targeting) und damit der Einschränkung der informationellen Selbstbestimmung reichen im Kontext von Big Data aber weiter und sind effizienter umsetzbar als bisher. Ein Problem liegt darin, dass man zwar grundsätzlich auf die Thematik aufmerksam machen kann; ob ein ethisches Problem aber wirklich vorliegt, muss im Einzelfall geprüft werden.

Kunden, die auf ihrem Recht der informationellen Selbstbestimmung beharren, könnten zudem das Problem haben, dass sie aufgrund der Verweigerung der Preisgabe von privaten Informationen eine Dienstleistung (z.B. eine Versicherung) nicht erhalten oder dafür mehr als andere bezahlen müssen (z.B. höhere Prämien für eine Versicherung). Wenn Unternehmen private Informationen benötigen, um eine Dienstleistung überhaupt anbieten zu können (z.B. Abklärung der Kreditwürdigkeit im Falle der Eröffnung eines Hypothekarkredits (Beispiel 1)) oder um ihr Risikomanagement zu verbessern (Beispiel 2), muss der Wert der informationellen Selbstbestimmung gegen die Ansprüche der Unternehmen abgewogen werden.

## Kontrolle der eigenen (digitalen) Identität

Eng mit der Forderung nach informationeller Selbstbestimmung verbunden ist die Forderung nach Kontrolle über die eigene digitale Identität. Es handelt sich um einen wichtigen Spezialfall der informationellen Selbstbestimmung. Digitale Identitäten können konstruiert werden, indem Unternehmen aufgrund von Big-Data-Anwendungen die Möglichkeit haben, verschiedene Merkmale eines Kunden zu aggregieren, zu korrelieren und zu einer digitalen Identität zu verdichten. Für sich genommen haben einzelne Daten wie beispielsweise das Tippverhalten oder die Uhrzeit von Telefonanrufen nichts mit der finanziellen Vertrauenswürdigkeit einer Person zu tun. Werden viele solcher Einzeldaten kombiniert und zu einer digitalen Identität geformt, erlauben sie jedoch Rückschlüsse auf den zukünftigen Schuldner und dessen Umgang mit Finanzen (seine Bonität). Diese Vorsichtsmass-

nahme, um den Ausfall von Debitoren zu vermeiden (Beispiel 1), ist ethisch dann problematisch, wenn der Kunde davon nichts weiss (vgl. Abschnitt «Transparenz») und/oder keine Möglichkeit hat, allfällige Fehler in diesem Bild von ihm zu korrigieren. Dieses Problem tritt nicht zuletzt dann auf, wenn die Möglichkeit des Vergessens oder Verjährens nicht vorgesehen ist. Personenbezogene Situationen und Verhaltensweisen werden zu einem späteren Zeitpunkt oft anders interpretiert als zu einem früheren – man denke etwa an die verbreitete Nutzung sozialer Medien in der Adoleszenz, in der «Jugendsünden» ihre digitalen Spuren hinterlassen. Unternehmen, die derartige Veränderungen nicht berücksichtigen und keine technischen Massnahmen treffen, um automatisierte Urteile zu verifizieren, behandeln potenzielle Kunden unfair, da diese an der Interpretation der Daten nicht mitwirken können.

Ethisch nicht problematisch ist allerdings, dass digitale Identitäten immer selektiv und einseitig sind, also der gesamten Identität einer Person nicht gerecht werden. Personen selbst präsentieren sich zuweilen ganz bewusst in unterschiedlicher Weise auf unterschiedlichen digitalen Medien, was auch als Ausdruck der informationellen Selbstbestimmung anzusehen ist, sofern keine Betrugsabsichten damit verbunden sind. Mehr noch, es ist ja gerade im Sinne des Datenschutzes, dass ein Unternehmen nur diejenigen Daten einer Person kennt und verwendet, die im Hinblick auf den Verwendungszweck (z.B. die Vergabe von Krediten) relevant sind.

Ebenfalls kein grundsätzlich neues Problem liegt vor, wenn sich Individuen in Antizipation der auf Algorithmen basierenden Entscheidungen von Unternehmen anders verhalten, sich für anderes interessieren und sich mit anderen Personen befreunden, um beispielsweise ihre Kreditwürdigkeit zu beeinflussen oder bestimmte Güter und Dienstleistungen günstiger zu erhalten. Dass es hier sozusagen zu einer «Fremdbestimmung» der Identität kommt, gehört zur Art und Weise, wie Menschen sozial interagieren und sich Vorteile zu verschaffen versuchen. Problematisch ist jedoch, wenn die durch Big Data ermöglichte permanente kommerzielle Überwachung durch Unternehmen Menschen überhaupt keine Freiräume für nicht-strategisches Verhalten mehr lässt, sie also keine andere Möglichkeit mehr sehen, als den vermuteten, durch Algorithmen vollzogenen Erwartungen zu entsprechen.

## Transparenz

Spricht man bezüglich der Wirtschaft von Transparenz, ist in der Regel gemeint, dass Kunden, Geschäftspartnern oder Investoren diejenigen Informationen über den Zustand des Unternehmens, seine Geschäftsabläufe, Dienstleistungen und Produkte zur Verfügung stehen, die sie benötigen, um rationale Entscheidungen zu treffen. Angestrebt wird eine offene und relevante Kommunikation bezüglich der zur Debatte stehenden Entscheidungen. Nur so können die Akteure sich eine gut informierte Meinung bilden und freie Entscheidungen treffen – Transparenz ist also auch eine Voraussetzung für informierte Zustimmung zur Verwendung persönlicher Daten. Märkte können ohne Transparenz nicht funktionieren. Im

Falle von Big Data gehört zur Forderung nach Transparenz das Recht jeder Person zu wissen, wer welche Daten von ihr oder über sie zu welchem Zweck verarbeitet.

Ein erstes Problem besteht im Falle vieler Datenerhebungen (aber selbstverständlich nicht nur hier) darin, dass die Unternehmen die Einwilligung der Nutzer in umfangreichen, für den juristischen Laien vielfach nur schwer verständlichen und oft wenig transparenten Allgemeinen Geschäftsbedingungen (AGB) einholen. Natürlich sind die Kunden angewiesen, derartige AGB vor der Zustimmung sorgfältig zu lesen. Aber die mangelnde Verständlichkeit und der Umfang vieler AGB lässt eine informierte Zustimmung häufig gar nicht zu.

Für die Nutzer ist in der Regel auch nicht transparent, auf welche Daten Plattformen und Apps zugreifen, wie lange diese Daten gespeichert werden, was analysiert wird sowie an wen und zu welchen Konditionen die Daten weitergegeben werden. Wenn beispielsweise App-Hersteller die von den Apps gesammelten Daten an Drittunternehmen verkaufen, ohne die Nutzer darüber transparent zu informieren, stellt dies eine Verletzung des Prinzips dar, dass Daten nur bezüglich der aus den Umständen zu erwartenden Zwecken bearbeitet werden dürfen. Wo dies passiert, steht dahinter wohl die Befürchtung der Unternehmen, dass die Kunden ihre Dienste meiden würden, wären die Big-Data-Praktiken transparent.

Im Falle massgeschneiderter Angebotskonditionen (Beispiel 3) sind die der Preisberechnung zugrundeliegenden Algorithmen für die Konsumenten nicht einsehbar, da sie zumeist dem Geschäftsgeheimnis unterliegen. Auch sind die im Hinblick auf personalisierte Angebote gesammelten Daten bezüglich ihrer Qualität, Richtigkeit und Vollständigkeit nicht überprüfbar. Während durch den Online-Handel Preisvergleiche einfacher wurden, verschlechtert sich durch Big-Data-Methoden die Preistransparenz wieder. Es kommt hier also zu einer Spannung zwischen dem berechtigten Anspruch eines Unternehmens als Urheber bestimmter Algorithmen deren Design (vor Konkurrenten) zu schützen und dem berechtigten Anspruch von Kunden auf Transparenz. Diese Herausforderung zu meistern ist Aufgabe der Unternehmen. Im Sinne einer gesellschaftlichen Verantwortung von Unternehmen darf sie nicht auf Kosten der Kunden gelöst werden.

Der Einsatz von lernenden Verfahren verschärft dieses Problem. Um Debitorenausfälle zu verhindern (Beispiel 1) sammeln Unternehmen Daten, um darin Korrelationen zu erkennen. Dabei ist für den Kunden in der Regel nicht transparent, welche Daten über ihn vorliegen und ob alle ihm zugeschriebenen Merkmale richtig zugeordnet, korrekt und im Hinblick auf den Zweck vollständig und relevant sind. Selbstlernende Algorithmen lernen mit jedem Geschäftsvorgang und mit jedem Debitorenausfall dazu und treffen eigenständig Entscheidungen, die selbst für die Verantwortlichen in den Unternehmen oft nicht mehr nachvollziehbar sind. Gerade neuere Ansätze wie das so genannte «deep learning»

sind damit eine grundsätzliche Herausforderung für das Gebot der Transparenz, weil auch die Systementwickler selbst nicht mehr wissen, wie das System zu seinen Schlüssen kommt. Statistische Korrelationen werden auf der Basis der Gesamtheit aller Kunden generiert und auf den individuellen Kunden übertragen, ohne dass das über alle feststellbare Muster für jedes Individuum zutreffend sein muss. In diesem Fall ist erneut der Wert der Autonomie betroffen, indem die Kontrolle der eigenen digitalen Identität eingeschränkt wird (vgl. Abschnitt «Kontrolle der eigenen digitalen Identität»).

## Solidarität

Solidarität bezeichnet die Verbundenheit der Einzelpersonen in einer Gemeinschaft. Im hier relevanten Sinn werden Einzelpersonen in Bezug auf bestimmte Risiken und Widerfahrnisse, denen sie in ihren Aktivitäten und Lebensvollzügen ausgesetzt sind, nicht allein gelassen. Die Gemeinschaft unterstützt sie (meist finanziell) bei der Bewältigung von Notsituationen wie Krankheiten und Unfällen sowie Armut. Dahinter steht die Einsicht, dass jeder ohne sein Verschulden z.B. krank oder arm werden kann.

Theoretisch könnten Versicherungen im Hinblick auf die Minimierung ihrer Risiken (Beispiel 2) Personen mit einer genetischen Prädisposition für gewisse Krankheiten oder Personen, die rauchen, sich ungesund ernähren oder zu wenig bewegen, vom Versicherungsschutz ausschliessen oder mit höheren Prämien belasten. Ein gängiges Kriterium, inwieweit eine Person Solidarität beanspruchen darf, ist dabei, ob die erwähnten Widerfahrnisse selbstverschuldet sind oder nicht (Verursacherprinzip). Die Kombination verschiedener Daten über das persönliche Verhalten (z.B. von Onlinekäufen, TV-Gewohnheiten, Bewegungsprofilen oder der Art der Nutzung sozialer Medien) kann mit dem Ziel durchgeführt werden, Risiken individualisiert zu beurteilen und als Ausdruck eines «selbstbestimmt gewählten Verhaltens» zu klassifizieren. Dies könnte einen ethisch legitimierten Grund zum Ausschluss grosser Risiken generieren. Solidarität würde damit durch die Hintertüre unterminiert, indem die Legitimität ihrer Inanspruchnahme hinterfragt würde. Oder aber man definiert mit Bezugnahme auf Solidarität Anreize zur Verhaltenssteuerung, um die genannten Risiken zu minimieren. So könnten Versicherungen von ihren Kunden z.B. die Einhaltung bestimmter Ernährungs- und Fitnessgewohnheiten verlangen und diese mithilfe am oder im Körper getragener Sensoren überprüfen. Diese für die Versicherungen ökonomisch attraktive Verhaltensnormierung steht in Konflikt mit dem Recht auf Selbstbestimmung.

Hebeln also z.B. Krankenversicherungen, die gesündere und gesünder lebende Versicherte begünstigen, den Solidaritätsgedanken als Grundprinzip des Versicherns aus? Das Verursacherprinzip könnte hier dazu eingesetzt werden, um die Solidarität der Versicherten einzugrenzen. Das Hervorheben des Verursacherprinzips ist allerdings dann problematisch, wenn bestimmte Krankheiten ausschliesslich bestimmten (letztlich frei gewählten) Verhaltensweisen zugeschrieben

und genetische, soziale und umweltbedingte Faktoren, für welche die Betroffenen nichts können, ausser Acht gelassen werden.

In der Schweiz dürfen Versicherer gemäss Bundesgesetz über die Krankenversicherung (KVG) Einwohner allerdings weder von der obligatorischen Grundversicherung ausschliessen noch deren Versicherungsprämie nach Einkommen oder Krankheitsrisiko differenzieren. Diese Bestimmungen sollen die Solidarität zwischen Versicherten mit unterschiedlichem Krankheitsrisiko stärken. Anders ist die Gesetzeslage bei den freiwilligen Versicherungen, die Angebote und Konditionen frei gestalten können. Da die relevanten Krankheiten mit der Grundversicherung in der Regel abgedeckt und die Zusatzversicherungen freiwillig sind, stellt die Aushebelung des Solidaritätsprinzips im Falle der Zusatzversicherungen kein grundsätzliches ethisches Problem dar.

## Kontextuelle Integrität

Die Lebenswelt des Menschen ist in verschiedene Bereiche gegliedert, die für das Individuum einen zentralen Orientierungsmassstab bilden. Menschen erwarten, in einem familiären Kontext anders behandelt zu werden als gegenüber einer staatlichen Organisation. Sie akzeptieren in einem ökonomischen Kontext Formen der Ungleichbehandlung, die man im Gesundheits-, Rechts- oder Bildungswesen nicht akzeptieren würde. Die Interpretation moralischer Grundwerte wie z.B. Gerechtigkeit, Autonomie und die damit verbundenen Regeln – im Fall von Gerechtigkeit Allokationsregeln wie «jedem das Gleiche» oder «jedem das, was er verdient oder braucht» – unterscheiden sich je nach sozialer Sphäre. Entsprechend unterscheiden sich auch die Informationen, die in diesen verschiedenen Sphären erzeugt und von Individuen preisgegeben werden – man spricht von der kontextuellen Integrität der Information. Wenn eine Person beispielsweise im Kontext des Gesundheitswesens persönliche Daten der medizinischen Forschung zur Verfügung stellt, ist der Wunsch, Dritten zu helfen, oft das entscheidende Motiv. Werden diese Informationen nun aber verwendet, um etwa Versicherungsangebote masszuschneidern (Beispiel 2) und Profite zu maximieren (Beispiel 5), so entspricht dies nicht mehr der ursprünglichen Absicht – die kontextuelle Integrität der Information wird verletzt.

Big-Data-Anwendungen, die danach streben, möglichst viele verschiedene Informationen über Individuen zu erfassen, tragen das inhärente Risiko in sich, diese kontextuelle Integrität zu verletzen. Da Daten zunehmend von Daten-Brokern gehandelt werden und in komplexe statistische Modelle von Personengruppen einfliessen, ist eine solche Verletzung der kontextuellen Integrität selbst für die kommerziellen Nutzer der Daten schwer erkennbar. Dies unterminiert damit auch das Gebot der Transparenz.

## Eigentums- und Urheberrecht

Das Wirtschaftssystem ist von ethischen Voraussetzungen abhängig, die es für sein Funktionieren braucht. Dazu gehören das Eigentums- und das Urheberrecht, die in der Schweiz von der Verfassung geschützte Grundrechte sind. Im Fall von Big Data stellt sich die Frage, inwieweit Daten ebenfalls unter diese Rechtsnormen fallen.

Die Nutzung von Diensten setzt vielfach die Herausgabe von persönlichen Daten voraus und der Datenhandel stellt für Unternehmen zunehmend eine zentrale Umsatz- und Innovationsquelle dar (Beispiel 5). Der Wert vieler Unternehmen bemisst sich nach der Anzahl ihrer Kunden und vor allem deren Daten. Dies wirft Fragen bezüglich des Eigentums an den Daten auf. Unbestritten ist, dass Kunden und Nutzer von Internet, Smartphone, Tablet und Smartwatch die Rohdaten generieren. Aber handelt es sich bei diesen Datenströmen um Werke, die im Sinne des Urheberrechts zu schützen sind? Unbestritten ist, dass die Daten geschäftlich erst dann nutzbar sind, wenn Unternehmen in die Rohdaten investieren und sie aufgrund ihres Wissens und ihrer Technologien bearbeitet haben.

Die offenen Fragen bezüglich der Eigentumsverhältnisse stellen die Geschäftsmodelle zahlreicher Unternehmen tiefgreifend in Frage. So beispielsweise, wenn sie ihre Nutzer an dem durch den Datennutzung und -handel erwirtschafteten Ertrag beteiligen müssten oder wenn Personen das Recht eingeräumt wird, dass ein Unternehmen Daten über eine Person löschen muss, wie dies die neue Datenschutz-Grundverordnung der EU vorsieht.

Aus ethischer Perspektive geht es darum, einerseits die Datengeber für die Zurverfügungstellung ihrer Daten und andererseits die Unternehmen für ihre Investitionen in diese Daten angemessen zu entschädigen. Für Letzteres ist im Rahmen unseres Wirtschaftssystems gesorgt: Unternehmen werden nur investieren, wenn sie ansprechende Gewinne erwarten können. Sie sind frei, das zu tun oder zu unterlassen. Anders sieht es bei den Rohdatenlieferanten aus: Ein gängiges Modell besteht darin, dass die Kunden für die Zurverfügungstellung der Daten durch die Gratisnutzung des Dienstes entschädigt werden. Hier sind zukünftig neue Modelle gefragt, z.B. die monetäre oder nicht-monetäre Partizipation der Rohdatenlieferanten am Erfolg der Dienste.

Inwieweit die Verletzung von kontextueller Integrität ein ethisches Problem darstellt, ist nicht einfach zu beurteilen, zumal die Grenzen der verschiedenen sozialen Sphären und der darin herrschenden moralischen Normen nicht unveränderlich sind. Die Gewichtung von Werten kann sich verschieben, wenn etwa die Individuen bereit sind, mehr Informationen aus ihrem privaten Bereich preiszugeben, und dies ebenso von ihren Mitmenschen erwarten. In diesem Fall verändert sich die geltende Vorstellung von Privatsphäre. Allerdings bleibt zu bedenken, dass die Ordnung der Welt in verschiedene soziale Sphären zentrale Bezugspunkte bildet. Dies erklärt, warum viele Menschen entrüstet reagieren würden, wenn z.B. private Informationen aus dem Freundeskreis zur Bildung individualisierter Preise (Beispiel 3) genutzt würden.

# Handlungsempfehlungen

Elektronisch erfasste und mit Algorithmen bearbeitete Daten durchdringen zunehmend unsere Gesellschaft und werden für die Entwicklung, Produktion und den Verkauf von Produkten und Dienstleistungen immer bedeutender. Die Ausführungen dieses Berichts thematisieren die Möglichkeiten, die sich Unternehmen durch das gezielte Erfassen und Analysieren von Daten eröffnen. Ein wichtiger Nutzen von Big Data liegt in der Individualisierung der Kunden, die aufgrund der immer zahlreicheren Sensoren immer mehr Datenspuren hinterlassen. Diese Daten ermöglichen es Unternehmen, den einzelnen Kunden immer treffsicherer z.B. hinsichtlich Verhalten und Vorlieben, Zahlungsmoral und Zahlungsbereitschaft, Kreditwürdigkeit und Risikogruppe einzuschätzen. Dieses Wissen erlaubt es Unternehmen, ihren Kunden individualisierte und dadurch relevante Angebote zu unterbreiten. Daten erhöhen die Effektivität der Angebote und die Effizienz der Werbung. Darüber hinaus stellen sie für Unternehmen eine immer wichtigere Innovations- und Umsatzquelle dar.

Big Data birgt einen Mehrwert für Unternehmen wie auch für die Kunden – für Letztere beispielsweise in Form von relevanteren und innovativen Angeboten und fokussierter Werbung. Je mehr Daten Unternehmen über ihre Kunden erheben können, desto präziser können sie diese ansprechen. Viele Angebote sind darüber hinaus in ihrer Funktionalität zwingend an die Preisgabe von Daten geknüpft.

Dass dabei die Big-Data-Anwendungen empfindlich in die Privatsphäre der Kunden eingreifen können, ist unbestritten und wird sowohl in der wissenschaftlichen Literatur als auch in den Medien breit diskutiert. Die ethische Debatte dieses Berichts zeigt, dass über den Schutz der Privatsphäre hinaus weitere ethische Werte und Normen von Big Data berührt werden, z.B. der Schutz vor Diskriminierung, Solidarität oder das Recht, die eigene (digitale) Identität zu kontrollieren. Dabei wird deutlich, dass Big-Data-Anwendungen nicht prinzipiell ethisch problematisch sind. Vielmehr müssen die Anwendungen im Einzelfall hinsichtlich ihrer Auswirkungen auf die betroffenen Anspruchsgruppen überprüft werden.

Schlussendlich geht es im ethischen Kontext von Big Data um ein Abwägen von Werten der betroffenen Anspruchsgruppen. Eingriffe z.B. in die Privatsphäre oder Ungleichbehandlungen von Kunden mit dem Ziel eines individualisierten Marketings und eines verbesserten Risiko- und Innovationsmanagements sind nachvollziehbare Interessen von Unternehmen. Diesen stehen die berechtigten Ansprüche der Kunden z.B. auf Schutz ihrer Privatsphäre und Nichtdiskriminierung gegenüber. Als ethisch problematisch erweist sich, wenn die Kunden nicht transparent über das Ausmass einer Datenerhebung und -nutzung informiert werden, sie für den Zweck einer bestimmten Datenverwendung nicht ihre Zustimmung gegeben haben oder sie keine realistische Alternative erhalten, sollten sie einer Datenherausgabe nicht zustimmen. All den ethischen Werten, die durch Big-Data-Anwendungen berührt werden, gemeinsam ist im Kern somit das Recht auf informationelle Selbstbestimmung, Transparenz und kontextuelle Integrität.

# Handlungsempfehlungen für Unternehmen

**Mit der voranschreitenden «Datafizierung» von Produktion und Konsum gewinnt auch die ethische Debatte über Big-Data-Anwendungen an Gewicht. Unternehmen können es sich zukünftig immer weniger erlauben, Daten zu sammeln, auszuwerten und weiterzugeben und die Information darüber im Kleingedruckten zu verstecken. Ihre Anspruchsgruppen werden informierter, kritischer und fordern zunehmend eine transparente Datennutzung. Letztendlich sind es die Kunden, die dem Unternehmen die Datenbasis liefern – je mehr Daten zur Verfügung stehen, desto besser und präziser werden die Big-Data-Anwendungen. Hierfür ist die gesellschaftliche Akzeptanz von entscheidender Bedeutung («licence to operate»). Im Rahmen ihrer geschäftlichen Aktivitäten müssen Unternehmen daher bestrebt sein, einen verantwortungsvollen Umgang mit Daten zu pflegen. Ein vertrauenswürdiger Datensammler zu sein und als solcher gesehen zu werden, wird zu einem immer wichtigeren Wettbewerbsvorteil.**

**Konkret empfiehlt es sich, dass Unternehmen für einen ethisch verantwortungsvollen Umgang mit Daten folgende Massnahmen treffen:**

**Ethics Case berücksichtigen**

Unternehmen sollten bei der Entwicklung von neuen Big-Data-Anwendungen neben dem Business Case von Anfang an auch die ethischen Aspekte berücksichtigen. Diese könnten z.B. im Rahmen bestehender Corporate Responsibility-Strukturen für jede Anwendung unternehmensintern eruiert werden. Dabei würde der Frage nachgegangen, welche ethischen Werte durch eine spezifische Big-Data-Anwendung berührt werden und ob man für die mit der spezifischen Anwendung verbundenen Zielsetzungen mögliche Konflikte mit den ethischen Werten und Interessen der Kunden riskieren will. Gleichzeitig sollte untersucht werden, ob die Erkenntnisse, die durch Big Data anhand vieler Datenpunkte generiert werden, wirklich effektiver sind als herkömmliche Methoden, die nur wenige Datenpunkte als Grundlage nutzen. Da die ethischen Aspekte in einzelnen Branchen dieselben sind, könnten auch branchenspezifische Agenturen oder Prüfstellen derartige Fragen aufnehmen und Branchenregelungen im Sinne einer Selbstverpflichtung der Wirtschaft vorschlagen.

**Kundenbedürfnisse evaluieren**

Die möglichen Folgen von Big-Data-Anwendungen auf die Unternehmensreputation und die Bereitschaft der Kunden, auch bei ausreichender Transparenz ihre Daten zur Verfügung zu stellen, können als Ethik-Indikator angesehen werden: Würden die Kunden ihre Daten immer noch zur Verfügung stellen, wenn sie wüssten, was mit ihren Daten geschieht? Um diese Frage zu beantworten, ist es für Unternehmen unabdingbar zu wissen, was der Kenntnisstand ihrer Kunden bezüglich Big Data ist, was sie von einem Unternehmen erwarten, welche Anwendungen sie akzeptieren und welche sie ablehnen und welche Nutzen für sie besonders relevant sind.

**Transparenz und Wahlfreiheit schaffen**
Das Vertrauen und die Akzeptanz der Nutzer als Datenlieferanten sind zwingende Voraussetzungen für einen erfolgreichen Einsatz von Big Data. Sie können nachhaltig nur sichergestellt werden, wenn Unternehmen proaktiv, transparent und in einer für die Kunden verständlichen Sprache über die Datenerhebung und -nutzung informieren. Unternehmen sollten beispielsweise über opt-in-basierte Lösungen die Zustimmung der Kunden zur Datenfreigabe einholen und ihnen angemessene Alternativen bieten, sollten sie einer Datenerhebung nicht zustimmen. Es liegt in der Verantwortung der Unternehmen, ihren Kunden eine angemessene Gegenleistung für die Daten zu bieten und ihnen diesen Mehrwert zu kommunizieren.

## Handlungsempfehlungen für die Politik

**Die ethische Debatte dieses Berichts macht deutlich, dass Big-Data-Ethik eine Frage des Abwägens ist, bei der für jeden Einzelfall geklärt werden muss, ob die Eingriffe in die ethischen Werte gerechtfertigt sind. Dabei ist nicht auszuschliessen, dass viele Unternehmen überfordert sein könnten, ohne von aussen gesetzte Leitplanken die richtige Balance zwischen kommerziellen Interessen und ethischen Prinzipien zu finden. Dies dürfte insbesondere der Fall sein, wenn Daten nicht nur für die Wertschöpfung im eigenen Unternehmen, sondern zusätzlich (oder gar primär) zum Verkauf an andere interessierte Unternehmen erhoben werden und ihre weitere Verwendung kaum mehr zu verfolgen ist.**

**Überarbeitung des Datenschutzgesetzes**
Die Ausführungen in diesem Bericht zeigen, dass der heutige Umgang mit Daten häufig im Gegensatz zu den Grundprinzipien des Datenschutzes steht. Gerade in der Akkumulation möglichst vieler Daten («V» für volume), in der Kombination verschiedenartiger Daten («V» für variety) und in ihrer Verwendung zu anderen Zwecken als denjenigen, zu denen sie ursprünglich erhoben wurden, liegt die Quintessenz von Big Data. Eine weitere, grundlegende datenschutzrechtliche Problematik von Big Data liegt darin, dass durch die grossen Datenvolumen und die Datenheterogenität auch bei anonymisierten Datensätzen oft Personenbezüge hergestellt werden können. Beim Datenschutzgesetz besteht daher Anpassungsbedarf. Die Revision der rechtlichen Rahmenbedingungen sollte so erfolgen, dass diese Innovationen im Bereich Big Data begünstigen und die Vorteile ausgeschöpft werden können, die die Technologie bietet. Gleichzeitig sollen die ethischen Überlegungen eine wichtige Rolle spielen, die in diesem Bericht dargelegt wurden.

**Kooperation von Staat und Wirtschaft**

Das Datenschutzgesetz kann mit den technologischen Entwicklungen kaum Schritt halten. Deshalb werden Unternehmens- oder Branchenkodizes (Code of Conduct), auf die sich Unternehmen selbst verpflichten, an Bedeutung gewinnen. Unternehmens- und Branchenregelungen zeigen zwar Schwächen bezüglich ihrer Durchsetzbarkeit und der Sanktionierung von Verstössen, die obigen Ausführungen machen aber deutlich, dass ein verantwortungsvoller Umgang mit Daten im Eigeninteresse der Unternehmen liegt, da nur dadurch die langfristige gesellschaftliche Akzeptanz der Big-Data-Anwendungen sichergestellt werden kann. In Anbetracht einer möglichen Überforderung der Unternehmen bei der ethischen Abwägung ist es Aufgabe des Staates, die Wirtschaft bei der Formulierung und Umsetzung von Verhaltenskodizes und Branchenregelungen zu unterstützen (z.B. im Rahmen von Round Tables oder durch die Förderung von Forschungs- und Entwicklungskooperationen). Denkbar ist darüber hinaus, dass eine staatlich akkreditierte, aber privat finanzierte Prüf- und Zertifizierungsstelle die ethische Abwägung vornimmt und die Einhaltung der Verhaltenskodizes überprüft. Diese Stelle würde prüfen, welche ethischen «Nebenwirkungen» eines Big-Data-Verfahrens vorliegen und ob der Zusatznutzen, den die Anwendung stiftet, Eingriffe z.B. in die Privatsphäre rechtfertigt.

**Standardisierung der AGB**

Eine transparente Nutzung der Daten ist die Grundvoraussetzung für die Wahrung der informationellen Selbstbestimmung der Kunden. Nutzer sollten proaktiv und transparent darüber informiert werden, wer welche Daten zu welchem Zweck über sie erhebt. Zu umfangreiche und für die meisten Nutzer schwer verständliche Allgemeine Geschäftsbedingungen erschweren oder verunmöglichen den Zugang zu diesem Wissen und somit eine informierte Zustimmung der Nutzer. Mit einer gesetzlich verankerten Vereinheitlichung der AGB könnte diesem Problem begegnet werden. Vorgaben bezüglich Inhalt, Formulierung und Form der AGB würden Transparenz schaffen und die Nutzer so befähigen, Ausmass und Art der Datenerhebung zu verstehen und eine informierte Entscheidung zu treffen.

# Glossar

**AGB** Allgemeine Geschäftsbedingungen

**App** Anwendungsprogramm, insbesondere für mobile Endgeräte wie Smartphones und Tablet-Computer

**Cookie** Ein Cookie (genauer: Browser-Cookie) ist ein Datensatz, den ein Server auf dem Endgerät des zugreifenden Nutzers ablegt, wenn dieser mit seinem Browser eine Website besucht. Bei einem erneuten Zugriff auf die gleiche Website kann der Server den Benutzer dann als den früheren Besucher wiedererkennen. Wenn eine andere Person den gleichen Computer benutzt, kann der Server dies allerdings nicht unterscheiden.

**DS-GVO** Datenschutz-Grundverordnung der EU. Gilt ab 25. Mai 2018 in allen EU-Mitgliedstaaten.

**DSG** Datenschutzgesetz (Schweiz)

**IKT** Informations- und Kommunikationstechnologie. Die Bezeichnung wird meist auf die elektronischen und heute weitgehend digitalen Technologien der Informationsverarbeitung und Telekommunikation bezogen.

**IP-Adresse** Eine von einem Computernetzwerk an den teilnehmenden Computer zugewiesene Adresse, die diesen im Netzwerk identifiziert. Aus der IP-Adresse, von der aus ein Nutzer z.B. auf eine Website zugreift, kann sehr grob auf seinen aktuellen Aufenthaltsort geschlossen werden.

**RFID** Radio Frequency Identification. Eine Technik zur kontaktlosen Identifikation durch Datenaustausch über Funkwellen, in der Regel über Entfernungen von unter einem Meter. Aus dem Alltag bekannte Anwendungen sind die «Funketiketten» in Verpackungen und Bibliotheksbüchern, RFID-Schlüssel, kontaktlos auslesbare Ausweise, Bank- und Kreditkarten sowie Chips zur Tieridentifikation.

**SCHUFA** Schutzgemeinschaft für allgemeine Kreditsicherung (Deutschland)

**Scopus** Ein vom Wissenschaftsverlag Elsevier betriebener Service zum Zugriff auf wissenschaftliche Publikationen, Patente und Zitate über das Internet.

**Scoring** Beim Scoring von Personen wird versucht, aus den über eine Person bekannten Daten deren zukünftiges Verhalten zu prognostizieren. Insbesondere wird beim Kreditscoring die Wahrscheinlichkeit ermittelt, mit der eine Person einer Zahlungsverpflichtung nachkommen bzw. nicht nachkommen wird.

**SIM** Subscriber Identity Module, auch «SIM-Karte» genannt, dient zur eindeutigen Identifikation eines Mobiltelefon-Nutzers durch die Stellen, mit denen das Mobiltelefon Daten austauscht.

**WoS** Web of Science. Ein vom Medienunternehmen Thomson Reuters betriebener Service zum Zugriff auf wissenschaftliche Publikationen über das Internet.

**ZEK** Zentralstelle für Kreditinformation (Schweiz)

# Weiterführende Literatur


Baeriswyl, B., Pärli, K. (Hrsg.) (2015):
**Datenschutzgesetz (DSG), Handkommentar.**
Bern: Stämpfli-Verlag.

BITKOM (2015):
**Big Data und Geschäftsmodell - Innovationen in der Praxis: 40+ Beispiele.**
Berlin: Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V.

Boyd, D., & Crawford, K. (2012):
**Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon.**
Information, Communication & Society, 15(5), 662-679.

Davis, K., & Patterson, D. (2012):
**Ethics of big data.**
Sebastopol, CA: O'Reilly Media.

Hilty, L. M.; Oertel, B.; Wölk, M.; Pärli, K. (2012):
**Lokalisiert und Identifiziert. Wie Ortungstechnologien unser Leben verändern.**
Zürich: vdf Verlag.

Hofstetter, Y. (2014):
**Sie wissen alles. Wie intelligente Maschinen in unser Leben eindringen und warum wir für unsere Freiheit kämpfen müssen.**
München: Bertelsmann.

Jarchow, T.; Estermann, B. (2015):
**Big Data, Chancen, Risiken und Handlungsbedarf des Bundes.**
Ergebnisse einer Studie im Auftrag des Bundesamts für Kommunikation.

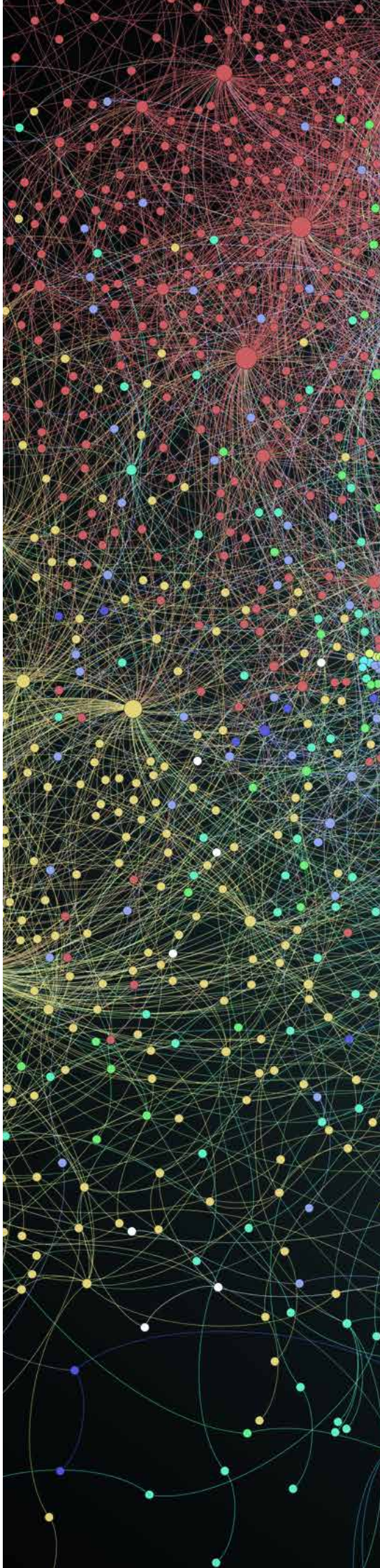
Maurer-Lambrou, U., Blechta, G.-P. (Hrsg.) (2014):
**Datenschutzgesetz, Öffentlichkeitsgesetz, Basler Kommentar (3. Auflage).**
Basel: Helbing Lichtenhahn Verlag.

Mayer-Schönberger, V., & Cukier, K. (2013):
**Big Data: die Revolution, die unser Leben verändern wird.**
München: Redline.

Passadelis, N., Rosenthal, D., Thür HP. (Hrsg.) (2015):
**Datenschutzrecht, Beraten in Privatwirtschaft und öffentlicher Verwaltung.**
Basel: Helbing Lichtenhahn Verlag.

Schneider, B. (2015):
**Data und Goliath: Die Schlacht um die Kontrolle unserer Welt: wie wir uns gegen Überwachung, Zensur und Datenklau wehren können.**
München: Redline.

Weber, R.H., Thouvenin, F. (Hrsg.) (2014):
**Big Data und Datenschutz – Gegenseitige Herausforderungen.**
Zürich: Schulthess Verlag.



satw it's all about technology

Schweizerische Akademie der Technischen Wissenschaften SATW
Gerbergasse 5 | 8001 Zürich | 044 226 50 11 | info@satw.ch | www.satw.ch