

**»Die zentrale Herausforderung auf dem Weg in eine ›data driven economy‹ ist, Deutschland zu einem international führenden Standort der Datenwirtschaft weiterzuentwickeln.«**

Thorsten Dirks  
Präsident, Bitkom e. V.

## **Leitlinien für den Big-Data-Einsatz**

Chancen und Verantwortung

### Herausgeber

BITKOM

Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e. V.

Albrechtstraße 10 | 10117 Berlin

Tel.: 030 27576-0 | Fax: 030 27576-400

bitkom@bitkom.org

www.bitkom.org

### Ansprechpartner

Dr. Mathias Weber | BITKOM e. V.

Tel.: 030 27576-121 | m.weber@bitkom.org

### Verantwortliches Bitkom-Gremium

AK Big Data in Zusammenarbeit mit AK Datenschutz

### Projektleitung

Arnd Böken, Graf von Westphalen Rechtsanwälte Partnerschaft

### Copyright

Bitkom, September 2015

Diese Publikation stellt eine allgemeine unverbindliche Information dar. Die Inhalte spiegeln die Auffassung im Bitkom zum Zeitpunkt der Veröffentlichung wider. Obwohl die Informationen mit größtmöglicher Sorgfalt erstellt wurden, besteht kein Anspruch auf sachliche Richtigkeit, Vollständigkeit und/oder Aktualität, insbesondere kann diese Publikation nicht den besonderen Umständen des Einzelfalles Rechnung tragen. Eine Verwendung liegt daher in der eigenen Verantwortung des Lesers. Jegliche Haftung wird ausgeschlossen. Alle Rechte, auch der auszugsweisen Vervielfältigung, liegen beim Bitkom.

# Leitlinien für den Big-Data-Einsatz

Chancen und Verantwortung

# Inhaltsverzeichnis

<b>1</b>	<b>Geleitwort</b>	<b>7</b>
<b>2</b>	<b>Geleitwort</b>	<b>11</b>
<b>3</b>	<b>Management Summary</b>	<b>14</b>
<b>4</b>	<b>Nutzen</b>	<b>22</b>
4.1	Energiewende	23
4.2	Verkehrssysteme	24
4.3	Medizinische Forschung und Diagnostik	26
4.4	Öffentlicher Bereich	31
4.5	Vorhersage von Krisensituationen	33
4.6	Bildung und Qualifizierung	35
4.7	Finanzdienstleistungen	37
4.8	Landwirtschaft	38
4.9	Connected Car – aus Sicht privater Nutzer	39
4.10	Digitale Einkaufswelten	40
4.11	Online-Marketing und Nutzen von Big Data für Verbraucher	41
4.12	Industrie 4.0	42
4.13	Telekommunikation	44
4.14	Verbesserung des Datenschutzes in Rechenzentren	45
<b>5</b>	<b>Bisher ungenutzte Chancen beim Einsatz von Big Data</b>	<b>50</b>
<b>6</b>	<b>Bedeutung der Datenwirtschaft in der zukünftigen Wirtschaft</b>	<b>56</b>
<b>7</b>	<b>Fragestellungen für Big-Data-Leitlinien</b>	<b>62</b>
7.1	Big Data im internationalen Wettbewerb	63
7.2	Vertrauensentwicklung bei Verbrauchern – Best Practices	65
7.3	Transparenz bei Big-Data-Analysen	66
7.4	Wie Analyseergebnisse zustande kommen	68
7.5	Bedeutung ethischer Grundsätze bei Big-Data-Lösungen	70
7.6	Stimulierung gesellschaftlich wünschenswerter Verhaltensweisen	72
7.7	Datensammlung im Automatismus	73
7.8	Recht auf Vergessenwerden?	74
7.9	Umgang mit Wahrscheinlichkeiten und Verhaltensprognosen	77
7.10	Erkenntnisse aus der Kommunikation in Sozialen Netzwerken	78
<b>8</b>	<b>Leitlinien für den Big-Data-Einsatz</b>	<b>82</b>
<b>9</b>	<b>Big Data und Datenschutz: Politischer Handlungsbedarf</b>	<b>88</b>
<b>10</b>	<b>Autorenverzeichnis</b>	<b>92</b>
<b>11</b>	<b>Verzeichnis der Abkürzungen</b>	<b>93</b>
<b>12</b>	<b>Sachwortverzeichnis</b>	<b>94</b>
<b>13</b>	<b>Quellen</b>	<b>97</b>

# Verzeichnis der Abbildungen

Abbildung 1: Big Data als Motor für Industrie 4.0	43
---	----

# Verzeichnis der Tabellen

Tabelle 1: Gesellschaftliche Implikationen autonomer Fahrzeuge	40
Tabelle 2: Einflussbereiche von Big Data	43

# Verzeichnis der Leitlinien

Leitlinie 1 – Nutzen der Big-Data-Anwendungen prüfen	82
Leitlinie 2 – Anwendungen transparent gestalten	82
Leitlinie 3 – Bevorzugt anonymisierte oder pseudonymisierte Daten verarbeiten	82
Leitlinie 4 – Interessen der Beteiligten abwägen	82
Leitlinie 5 – Einwilligungen transparent gestalten	82
Leitlinie 6 – Nutzen für Betroffene schaffen	82
Leitlinie 7 – Governance für personenbezogene Daten etablieren	82
Leitlinie 8 – Daten wirksam gegen unberechtigte Zugriffe schützen	82
Leitlinie 9 – Keine Daten zu ethisch-moralisch unlauteren Zwecken verarbeiten	82
Leitlinie 10 – Datenweitergabe nach Interessenabwägung ermöglichen	83
Leitlinie 11 – Selbstbestimmtes Handeln ermöglichen	83
Leitlinie 12 – Politische Rahmenbedingungen vervollkommen – Datenschutz und Datennutzen neu abwägen	83



# 1 Geleitwort

Epochale Chancen:  
Die digitale Transformation  
erfolgreich gestalten





# 1 Geleitwort

## Epochale Chancen: Die digitale Transformation erfolgreich gestalten

Wir erleben derzeit die spannendste Revolution, die es in der Wirtschafts- und Technologiegeschichte je gab. In den kommenden zehn Jahren wird sich die Welt so stark verändern wie nie zuvor. Dabei ist es nicht nur das Ausmaß der Veränderung, es ist ebenso ihre Geschwindigkeit, die das Hier und Jetzt so einmalig macht. Die industrielle Revolution dauerte über einhundert Jahre, fünf Generationen. Die digitale Revolution wird zehn, maximal zwanzig Jahre benötigen. Eine Generation, mehr nicht. Ihre Bedeutung für jeden Einzelnen von uns, für die Gesellschaft, unser Wertesystem, die Politik, wird dabei in nichts hinter dem zurückstehen, was vormals die industrielle Revolution mit sich brachte.

Eine erfolgreiche digitale Transformation wird zusätzlichen Wohlstand schaffen. Dabei fallen enorme Datenmengen an. Professor Roberto Zicari von der Goethe-Universität Frankfurt schätzt, dass jeden Tag etwa 2,5 Trillionen Bytes Daten neu generiert werden. Derzeit generieren wir alle zwei Tage so viele Daten wie vom Beginn der Menschheitsgeschichte bis zum Jahr 2003 – so eine Schätzung Eric Schmidts. Die digitale Revolution wird nur funktionieren, wenn wir diese enormen Mengen sinnvoll strukturieren und analysieren können und verantwortungsvoll mit ihnen umgehen.

Big Data Technologien, mit denen sich die exponentiell wachsenden Datenmengen zielorientiert und nahezu in Echtzeit analysiert werden können, sind inzwischen verfügbar. So lassen sich Krebs besser heilen, Herzinfarkte leichter vermeiden und der Straßenverkehr so steuern, dass es nicht nur weniger Staus und Schadstoffausstoß, sondern auch weniger Unfallopfer gibt.

Viele denken bei dem Begriff Big Data unwillkürlich an Big Brother, an eine Bedrohung orwellischen Ausmaßes. Für die Politik heißt es, bewusst und behutsam mit dem Thema umzugehen. Weder darf man Risiken ausblenden, noch darf man aus Angst vor Risiken die Chancen ungenutzt lassen. Dieses Positionspapier zeigt insbesondere die Chancen auf.

Wir sind auf dem Weg in eine „data driven economy“. Die zentrale Herausforderung ist, Deutschland zu einem international führenden Standort der Datenwirtschaft weiterzuentwickeln.

**Thorsten Dirks**

Präsident Bitkom e. V.



# 2 Geleitwort

Big-Data-Zeitalter:  
Herausforderungen erkennen  
und gestaltend annehmen



## 2 Geleitwort

### Big-Data-Zeitalter: Herausforderungen erkennen und gestaltend annehmen

Die Digitalisierung unserer Welt verändert unser Leben, revolutioniert ganze Industrien, eröffnet neue Geschäftsmodelle und stellt uns in Wirtschaft und Gesellschaft vor neue und große Herausforderungen. Sie ist in allen Lebens- und Unternehmensbereichen mit einem rapiden Wachstum an digitalen Datenbeständen verbunden, weshalb gerne auch vom aufkommenden Big-Data-Zeitalter gesprochen wird. Jedes Ding und jeder Vorgang der realen Welt bekommt eine digitale Hülle, über die sie mit anderen Dingen und Vorgängen IT-technisch verknüpft werden können, und vermittels derer Kommunikation und intelligente Interaktion möglich wird.

Der Umgang mit diesen großen Datenmengen und ihre Nutzung ist die Grundlage für die digitale Transformation und bestimmt Ausrichtung und Geschwindigkeit des sich vollziehenden Wandels. Ihm kommt deshalb eine grundlegende Bedeutung zu, nicht nur für die Wirtschaft, sondern für alle Felder des gesellschaftlichen Miteinanders. Wenn es uns in Deutschland nicht gelingt, diese Herausforderung zu erkennen und gestaltend anzunehmen, wird dies gravierende Auswirkungen auf unsere weitere wirtschaftliche und gesellschaftliche Entwicklung und den Wohlstand jedes Einzelnen haben; es besteht die Gefahr, digitale Kolonie zu werden.

Ein ganz wichtiger Aspekt beim Thema digitale Transformation und Big Data ist die Frage, welche Auswirkungen Big-Data-Techniken auf unser persönliches Leben haben werden. Deutschland beschwört eine besondere Tradition im Kontext des Datenschutzes; Prinzipien wie das Recht auf informationelle Selbstbestimmung oder das Konzept der Datensparsamkeit und Datenvermeidung waren bewährte Richtlinien in der Vergangenheit. Doch liefern diese Prinzipien auch die richtigen Antworten auf die Herausforderungen des neuen Big-Data-Zeitalters? Hier bedarf es dringend eines umfassenden gesellschaftlichen Dialogs, in dem die Tauglichkeit dieser Prinzipien für das Big-Data-Zeitalter hinterfragt und wo immer nötig kreativ weiterentwickelt werden. Dabei braucht es Antworten, die es erlauben, die sich bietenden Potenziale zu erschließen und gleichzeitig möglichen Fehlentwicklungen von vornherein einen Riegel vorzuschieben.

Um diesem Dialog eine verlässliche Basis zu geben, braucht es eine solide und vertrauenswürdige Informationsgrundlage, die auf der Basis von konkreten Pilotanwendungen verlässlich Auskunft über Chancen und Risiken der Big-Data-Technologie gibt und so eine Richtschnur aufzeigt zur Entwicklung ethischer Grundsätze für einen einwandfreien und konsensfähigen Einsatz von Big-Data-Analysen. Das vorliegende Positionspapier leistet hierzu einen wichtigen Beitrag.

**Prof. Dr. Christoph Meinel**

Institutsdirektor und CEO, Hasso-Plattner-Institut GmbH



# 3 Management Summary

## 3 Management Summary

### Ziele, Zielgruppen, Schwerpunkte

Diese Publikation adressiert Entscheidungsträger in der Politik und in den Datenschutzbehörden, private Verbraucher sowie die breite Öffentlichkeit und die Medien. Es wendet sich ebenfalls an die Nutzer von Big-Data-Technologien und -Lösungen in der Wirtschaft, deren Anbieter sowie Ethikbeauftragte. Sie verfolgt das Ziel, rechtspolitische Positionen und Leitlinien für den ethisch fundierten, verantwortungsbewussten Einsatz von Big Data zu formulieren. Die Publikation soll auch zu mehr Transparenz im Bereich Big-Data-Anwendungen beitragen, denn Information und Transparenz schaffen die Basis für Vertrauen.

Das Positionspapier entwickelt die Handlungsempfehlungen aus dem Strategiepapier für den Nationalen IT Gipfel 2014<sup>1</sup> weiter und konzentriert sich auf zwei Schwerpunkte:

- Mit Beispielanwendungen wird der vielfältige Nutzen für die Verbraucher und die Gesellschaft insgesamt gezeigt, den Big-Data-Anwendungen mit sich bringen. Die Beispiele stammen aus vielen Bereichen des gesellschaftlichen Lebens – u. a. aus den Einsatzgebieten Hochwasserschutz, Verkehrsplanung und -überwachung, Forensik, aus der medizinischen Forschung und Diagnostik sowie aus Wirtschaftszweigen, mit deren Produkten und Services private Verbraucher besonders oft konfrontiert werden.
- Den zweiten Schwerpunkt bilden Fragestellungen, die im Zusammenhang mit Big-Data-Anwendungen im Fokus des öffentlichen Interesses stehen. Dazu gehören u. a. der Datenschutz, die Vertrauensentwicklung, die Transparenz, die automatisierte Datensammlung sowie die ethische Fundierung (vgl. Kapitel 7). Aus diesen Darlegungen im Kapitel 7 werden dann Leitlinien<sup>2</sup> für den Big-Data-Einsatz (Kapitel 8) und Vorschläge an die Politik (Kapitel 9) abgeleitet.

### Nutzen von Big-Data-Anwendungen für die Verbraucher und die Gesellschaft

Big Data bezeichnet den Einsatz großer Datenmengen aus vielfältigen Quellen mit einer hohen Verarbeitungsgeschwindigkeit zur Erzeugung wirtschaftlichen Nutzens.<sup>3</sup> Big-Data-Technologien werden in sehr vielen Bereichen eingesetzt und können zu einem hohen Nutzen führen – für Unternehmen, für Verbraucher und die Gesellschaft<sup>4</sup> insgesamt (vgl. Kapitel 4). Einsatzbeispiele aus 14 Bereichen werden vorgestellt – hier nur eine Auswahl:

- 1 Das vorliegende Positionspapier knüpft insbesondere an die Handlungsempfehlung »Schaffung und Berücksichtigung rechtlicher und ethischer Rahmenbedingungen« an – vgl. (Projektgruppe Smart Data, 2014, S. 17)
- 2 Leitlinie wird hier als Handlungsempfehlung verstanden.
- 3 Vgl. (Bitkom, 2012). Bitkom hat eine Serie weiterer Entscheidungshilfen zu Big Data vorgelegt – vgl. (Bitkom, 2013), (Bitkom, 2014), (Bitkom, 2015) und (Bitkom, 2015a). Eine Übersicht über die Anbieterlandschaft bietet u. a. eine Untersuchung der Experton Group – vgl. (Landrock et al., 2014).
- 4 Prof. Roberto Zicari fördert am Big Data Lab der Goethe-Universität Frankfurt Ansätze, Big Data zum Nutzen der Gesellschaft einzusetzen, vgl. (Drilling/Litzel, 2015).



- Big-Data-Anwendungen im Gesundheitswesen unterstützen die Diagnose und Therapie von Krebserkrankungen, die Hilfe für Infarkt Betroffene, die Bereitstellung personalisierter Medikamente z. B. gegen Erbkrankheiten oder die Überwachung von Vitalparametern von Frühgeborenen, Senioren oder chronisch Kranken.
- Im öffentlichen Bereich werden Big-Data-Technologien zurzeit vorrangig für Sicherheitsaufgaben, die Bekämpfung von Terrorismus und organisierter Kriminalität sowie im Katastrophenmanagement eingesetzt. Andere europäische Länder nutzen Big-Data-Anwendungen für öffentliche Aufgaben, bei denen in Deutschland besondere Rahmenbedingungen gelten, die bisher dem Einsatz entgegenstanden. Viel wird davon abhängen, wie die in Diskussion befindliche EU-Datenschutz-Grundverordnung<sup>5</sup> konkret ausgestaltet wird.
- Sensor-, Cloud- und Big-Data-Technologien gehören zu den Grundlagen für autonome Fahrzeuge und lassen tiefgreifende Veränderungen bei Mobilitätslösungen erwarten. Diese Technologien leisten auch einen unverzichtbaren Beitrag für Industrie 4.0 sowie für intelligente Stromnetze und Verkehrssysteme der Zukunft.
- Mit Big-Data-Technologien erhöhen die Telekommunikationsunternehmen die Leistungsfähigkeit ihrer Infrastruktur. Sie können außerdem anonymisierte Bewegungsmuster ihrer Kunden erzeugen, die u. a. bei der Verkehrsplanung und -steuerung hilfreich sind.

Insgesamt wird die Datenwirtschaft die Geschäftsmodelle vieler Branchen unserer Industrie- und Dienstleistungslandschaft umkrempeln. Big Data wirkt als Katalysator für die Entstehung bzw. Modifizierung von Geschäftsmodellen<sup>6</sup>, ist bereits unternehmerische Realität und eröffnet neue Möglichkeiten für innovative Unternehmen und Startups.

Transparenz über den Nutzen von Big Data in vielen Einsatzgebieten hilft der Politik, informierte Entscheidungen zu treffen.

## Chancen ergreifen, Risiken eindämmen

Die vielen Einsatzmöglichkeiten dürfen nicht darüber hinwegtäuschen, dass Big Data nicht frei von Risiken ist, wenn es um personenbezogene Daten geht oder wenn Kausalzusammenhänge falsch interpretiert werden.<sup>7</sup>

Unsere Gesellschaft steht vor der Herausforderung, Missbrauch zu verhindern, Freiheitsrechte zu bewahren und die Privatsphäre zu sichern. Diesem Ziel dient der Dreiklang aus Gesetzen, Selbstverpflichtungen und ethischen Leitlinien der Wirtschaft und öffentlichem Diskurs in einer demokratisch verfassten Gesellschaft. Dieser Dreiklang stärkt die Immunität unserer Gesell-

<sup>5</sup> Vgl. <http://www.eu-datenschutzverordnung.de>

<sup>6</sup> Vgl. (Bitkom, 2015).

<sup>7</sup> Vgl. Kapitel 7 in (Bitkom, 2014).

schaft gegen ein Abgleiten in eine orwellsche Welt und bewahrt unsere demokratischen Grundsätze und Wertvorstellungen sowie das grundsätzliche Prinzip der Solidargemeinschaft.

Letztlich ist es ein gesellschaftlicher Aushandlungsprozess, wie mit den Ergebnissen bestimmter Big-Data-Analysen umgegangen wird. Klar ist: Für den Einsatz von Big-Data-Prognosen gibt es Grenzen. Prognosen dürfen nicht die Chancengleichheit von Menschen einschränken oder Menschen diskriminieren. Staatliche Zwangsmaßnahmen dürfen nicht an prognostizierten, sondern nur an tatsächlich begangenen Handlungen ansetzen. Die Kombination aus Gesetzen, in der Wirtschaft gelebten Werten und der informierten Öffentlichkeit wird verhindern, dass ambitionierte Big-Data-Anbieter bei der »Vermessung« des Menschen über das Ziel hinaus-schießen.

Chancen und Risiken müssen bei jeder Anwendung konkret abgewogen werden – Pauschalurteile genügen hier nicht. Neuen Technologien schlägt immer Skepsis und auch Ablehnung entgegen. In der breiten Öffentlichkeit ist zu wenig über erfolgreiche Big-Data-Anwendungen bekannt. Mehr Informationen, mehr Fallbeispiele, mehr Transparenz tragen zur Akzeptanz bei. Hier sind auch die Anbieter gefordert.

## Schwerpunkte im öffentlichen Diskurs

Der Einsatz von Big-Data-Technologien ist Gegenstand eines intensiven öffentlichen Diskurses. In Kapitel 7 werden ausgewählte Fragen aufgegriffen, die aus Sicht des Bitkom besonders wichtig sind.

- **Rechtssicherheit auf europäischer Ebene:** Für Big Data in der internationalen Dimension sind einheitliche, international geltende Regeln erforderlich. Die in Diskussion befindliche EU-Datenschutz-Grundverordnung bleibt derzeit noch hinter den Möglichkeiten und Erfordernissen der Digitalisierung zurück. Es ist Aufgabe der Politik, für die angemessenen Rahmenbedingungen zu sorgen.
- **Transparenz und Verbraucherakzeptanz von Big Data:** Damit die Verbraucher Big-Data-Anwendungen akzeptieren, müssen ihre Vorbehalte und Ängste verstanden und ernst genommen werden. Transparenz ist ein wesentlicher Schlüssel für die Akzeptanz von Big-Data-Lösungen in der öffentlichen Meinungsbildung. Die Technologieanbieter sollten sich zu einer ethisch fundierten Herangehensweise bekennen und Transparenz herstellen.
- **Bedeutung automatisierter Datensammlung:** Die automatisierte Sammlung von Daten bietet sowohl für den Verbraucher als auch für Unternehmen einen erheblichen Mehrwert. Zu gewährleisten ist dabei aber, dass die rechtlichen Vorgaben beachtet werden und der Verbraucher den Überblick darüber behält, welche Daten seine Geräte übermitteln, um selbstbestimmt handeln zu können.

- **Recht auf Vergessenwerden:** Die Forderung eines »Rechts auf Vergessenwerden« geht an der technologischen Realität vorbei. Vielmehr geht es um selbstbestimmtes Löschen und darum zu erfahren, wo welche personenbezogenen Daten liegen.
- **Grenzen für Prognosen:** Prognosen auf der Basis von Big-Data-Einsatz können an Grenzen geraten – sie dürfen nicht die Chancengleichheit von Menschen beschneiden.
- **Analyse Sozialer Netzwerke:** Unternehmen sollten bei der Analyse der Kommunikation in Sozialen Netzwerken umsichtig agieren und die Interessen der Betroffenen angemessen berücksichtigen.

## Leitlinien für den Big-Data-Einsatz

Da nicht sämtliche Einzelheiten des Wirtschaftslebens durch Gesetze geregelt werden können, sind Leitlinien für die Unternehmen als Orientierungsrahmen bei der Entwicklung und Nutzung von Big-Data-Technologien hilfreich.

Die Leitlinien sollen Hinweise für die Kommunikation mit Verbrauchern geben, Best Practices für das Prozessmanagement und die Unternehmensorganisation bieten und den Unternehmen insgesamt eine Richtschnur für einen ethisch und juristisch einwandfreien Einsatz von Big-Data-Analysen aufzeigen.

Besondere Verantwortung tragen Organisationen, wenn sie personenbezogene Daten verarbeiten. Dieses Positionspapier zeigt auch die Grenzen für den Big-Data-Einsatz auf. Mit einem entsprechenden Design von Prozessen und Organisation sowie mit Programmierungs-Know-how wappnen sich Unternehmen gegen Fehler im Umgang mit Big Data.

Die vorliegende Publikation ist Teil eines länger dauernden Meinungsbildungsprozesses in der Gesellschaft<sup>8</sup>, steht doch die Nutzung von Big Data in Wirtschaft und Verwaltung noch am Anfang. Der Diskurs über Chancen und Risiken kann nur anhand konkreter Anwendungen geführt werden. Demzufolge wird es erforderlich sein, die Leitlinien im Zuge der öffentlichen Diskussion weiterzuentwickeln. Dann wird auch erkennbar, ob die generischen Leitlinien in eine Selbstverpflichtung<sup>9</sup> der Big-Data-Anbieter münden können oder ob es weiterer Regulierung bedarf.

## Politischer Handlungsbedarf

Dieses Positionspapier ist ein Plädoyer für einen fortwährenden gesellschaftlichen und politischen Diskurs, um gesellschaftlich akzeptierte Standpunkte zu Big Data zu entwickeln. Chancen und Risiken müssen in diesem Diskurs in einen angemessenen Ausgleich gebracht werden und

---

8 Mit der »Schaffung und Berücksichtigung ethischer Rahmenbedingungen« befasst sich auch die Fachgruppe »Wirtschaftliche Potenziale und gesellschaftliche Akzeptanz«, die im Rahmen der Begleitforschung für das BMWi-Technologieprogramm »Smart Data – Innovationen aus Daten« etabliert wurde. Vgl. (BMW, 2015).

9 Vgl. (Swisscom, o.J.), (Deutsche Telekom, o.J.).

an diesem Maßstab sollten sich die rechtlichen Vorgaben orientieren. Bisher wurde Big Data lediglich unter Schutzaspekten gesetzlich erfasst. Es gilt jedoch, positive Nutzungsregeln ins Auge zu fassen, die der zentralen Bedeutung von Daten in einer Data Driven Economy gerecht werden.

## Struktur des Positionspapiers

Das Positionspapier umfasst sechs Kapitel, die durch Geleitworte, Management Summary sowie Verzeichnisse der Autoren, Sachwörter und Quellen ergänzt werden.

- Das Kapitel 4 soll zeigen, welchen Nutzen Big Data für die Gesellschaft insgesamt hat. Das gilt auch für eine so wichtige Entwicklung wie Industrie 4.0, die das Arbeitsleben zukünftig grundlegend umgestalten wird. Da die Einsatzgebiete und konkreten Anwendungen in der Öffentlichkeit noch zu wenig bekannt sind, legen die Autoren Wert auf Vielfalt und verständliche Darstellung. Für den schnellen Leser sind auch alle Abschnitte dieses Kapitels mit kurzen Zusammenfassungen versehen.
- Impulse für den Big-Data-Einsatz in Wirtschaft und Verwaltung kann auch der Blick über die Landesgrenzen hinaus vermitteln. Im Kapitel 5 werden Beispiele für Big-Data-Anwendungen vorgestellt, die aus verschiedenen Gründen in Deutschland (noch) nicht zum Tragen kommen.
- Das Kapitel 6 richtet den Blick über einzelne Big-Data-Anwendungen hinaus und charakterisiert die Datenwirtschaft als dynamische neue Branche der Volkswirtschaft, die Unternehmen vieler Branchen Geschäftsmodell-Innovationen und Wachstumspotenziale eröffnet.
- Das Kapitel 7 greift wichtige Aspekte aus dem öffentlichen Diskurs um den Big-Data-Einsatz auf und stellt Positionen des Bitkom vor.
- Kapitel 8 formuliert prägnante Leitlinien.
- Kapitel 9 wendet sich mit Vorschlägen an die Politik.





# 4 Nutzen von Big-Data- Anwendungen für Gesellschaft und Privatpersonen

## 4 Nutzen

### von Big-Data-Anwendungen für Gesellschaft und Privatpersonen

Das Kapitel 4 zeigt an Beispielen aus 14 ausgewählten Bereichen die Bedeutung von Big-Data-Technologien für die Gesellschaft und die Verbraucher auf.<sup>10</sup>

- Die durchgehende Digitalisierung der meisten Geschäftsprozesse in Unternehmen und Organisationen führt zu einem enormen Anstieg der Daten. Die dafür erforderliche Verbesserung von Dateninfrastrukturen steigert die Wettbewerbsfähigkeit der deutschen Wirtschaft und prägt maßgeblich die Attraktivität des Wirtschaftsstandorts Deutschland.<sup>11</sup> Von modernen Dateninfrastrukturen profitieren u. a. die Energieversorgung, das Verkehrswesen, das Gesundheitswesen und die staatliche Verwaltung. Wegen ihrer Bedeutung für die privaten Verbraucher werden auch Beispiele aus Bereichen wie Finanzdienstleistungen, Mobilität (Connected Car) oder Telekommunikation aufgeführt. Zusätzlich wird gezeigt, dass Industrie 4.0 aus Big-Data-Ansätzen maßgebliche Impulse erhält – auch wenn der Nutzen für private Verbraucher nicht unmittelbar ersichtlich wird.
- Die Bereitstellung von Nahrungsmitteln für eine auch weiter rasant steigende Weltbevölkerung ist eine eminent politische Aufgabe. Insofern sind die Möglichkeiten des Big-Data-Einsatzes in der Landwirtschaft von enormer gesellschaftlicher Relevanz.
- Das Internet der Dinge und Big Data dringen schrittweise von der Produktionskette bis an die Ladentheke. Big Data gibt dem (Online-)Händler Werkzeuge in die Hand, die potenziellen Käufer mit individuellen Angeboten zu adressieren und bis zur Kaufentscheidung zu begleiten.
- Big Data trägt zur Erhaltung eines kostenfreien, da werbefinanzierten Angebots-Ökosystems im Internet bei. Verbraucher profitieren von qualitativ hochwertigen und kostenfreien Inhalten, Dienstanbieter von optimierten Erträgen aus der Bereitstellung ihrer Services.

Die Datenwirtschaft wird die Geschäftsmodelle vieler Branchen unserer Industrie- und Dienstleistungslandschaft umkrempleln. Zahlreiche Beispiele aus Unternehmen zeigen die Entstehung bzw. Modifizierung von Geschäftsmodellen durch Big Data auf. Big Data ist unternehmerische Realität und eröffnet neue Möglichkeiten für innovative Unternehmen und Startups. Der Wert eines Produktes aus Verbrauchersicht wird immer stärker durch intelligente Software-Funktionalität und kontext-sensitive Bereitstellung von Daten geprägt (digitale Anreicherung).

<sup>10</sup> Zahlreiche Einsatzbeispiele von Big-Data in der privaten Wirtschaft enthalten die Leitfäden des Bitkom (Vgl. insbesondere (Bitkom 2012) und (Bitkom 2015)). Weitere wurden auf den Kongressen »Big Data Summit« vorgestellt (vgl. [www.bitkom-bigdata.de](http://www.bitkom-bigdata.de)).

<sup>11</sup> Mit dem Big-Data-Einsatz in Bereichen wie Energie- und Verkehrswesen werden enorme Ressourcen gespart und Beiträge zum Umweltschutz erbracht. Damit wird der Forderung entsprochen, dem Leitbild der nachhaltigen Entwicklung zum Durchbruch zu verhelfen. Vgl. dazu (Kreibich, 2015).



## 4.1 Energiewende

Das intelligente Stromnetz der Zukunft besteht aus einer großen Zahl von technischen Systemen, die durch Daten-Dienste miteinander verbunden sind. Dazu gehören die Erfassung, Speicherung, Verarbeitung und Visualisierung aller technischen und betriebswirtschaftlichen Daten und komplexe Optimierungsrechnungen. Big Data, Cloud Computing, Datenschutz und IT-Sicherheit sind für Smart Grids essenziell.

Die Bundesregierung plant den Einsatz von Smart Metering in deutschen Haushalten. Die Optimierung des privaten Verbrauchs von Elektroenergie leistet einen bedeutenden Beitrag zur Energiewende.

2013 hat das jetzige Bundesministerium für Wirtschaft und Energie eine Studie für den flächen-deckenden Einsatz von Smart Metering in Deutschland vorgestellt.<sup>12</sup> Ein Szenario geht davon aus, dass bis 2022 mehr als zehn Millionen digitale Zähler installiert sind und die Innovation somit rund zwei Drittel aller Haushalte und Unternehmen erfasst.

Smart Meter bilden die Grundlage für ein intelligentes Stromnetz der Zukunft (Smart Grid). Das intelligente Stromnetz verbindet Erzeuger und Verbraucher sowie Energiespeicher und ermöglicht den Austausch von Zustandsinformationen. So wird es möglich, Angebot und Nachfrage besser aufeinander abzustimmen und zeitnah auf Schwankungen zu reagieren. Das kann nur funktionieren, wenn es gelingt, die Daten von Millionen Verbrauchern in kurzer Zeit zu verarbeiten und eine Rückkopplung an die Verbraucher z. B. in Form von Preisinformationen zu gewährleisten, um die Motivation zur Verminderung des Verbrauchs in Zeiten von vergleichsweise geringem Angebot zu stärken.

Das intelligente Stromnetz der Zukunft besteht aus einer großen Zahl von technischen Systemen, die durch Daten-Dienste miteinander verbunden sind. Dazu gehören die Erfassung, Speicherung, Verarbeitung und Visualisierung aller technischen und betriebswirtschaftlichen Daten und komplexe Optimierungsrechnungen. Big Data, Cloud Computing und IT-Sicherheitstechnologien gehören zu den Technologien, die im intelligenten Stromnetz eingesetzt werden.<sup>13</sup>

Der Nutzen eines intelligenten Stromnetzes für Privatverbraucher ist offensichtlich: Sie erhalten eine deutlich höhere Transparenz über ihren Verbrauch und können Kostentreiber im Haushalt identifizieren und abstellen.

---

<sup>12</sup> Vgl. (Ernst & Young, 2013).

<sup>13</sup> Vgl. (Daniel, 2014) sowie (Bitkom, 2013, S. 88-90).

## 4.2 Verkehrssysteme

Der Einsatz von Big-Data-Lösungen im Luft-, Schienen- und Straßenverkehr ermöglicht jedes Jahr Millioneneinsparungen und verbessert die Umweltbilanz entscheidend. Gleichzeitig können die Beförderungsangebote besser auf den Bedarf der Passagiere angepasst werden.

### Optimierung im Schienen- und Straßenverkehr – am Beispiel der Schweiz

Aus der Logistik stammt eine Big-Data-Anwendung, die jedem Zugfahrer einen Nutzen klar vor Augen führt<sup>14</sup>: Die Schweizer Bahnen regeln die Geschwindigkeit ihrer Züge mit neuen Technologien. Die Verminderung von Verkehrshalten lässt auch den Energieaufwand sinken – mit Millioneinsparungen in jedem Jahr und Beiträgen zum Umweltschutz. Die Spezifik des Eisenbahnwesens führt dazu, dass Verbesserungen im System nur durch eine Optimierung des vorhandenen Netzes möglich sind. Das neue Rail Control System speichert und verarbeitet in Sekundenschnelle die Zuginformationen aller in Betrieb befindlichen Bahnen sowie Zustandsdaten der gesamten Bahninfrastruktur. Es liefert dem Lokführer Informationen zur Steuerung der Zuggeschwindigkeit mit dem Ziel, ungeplante Stopps zu vermeiden, denn starkes Abbremsen und nachfolgendes Beschleunigen großer Lasten kosten sehr viel Energie.

In ähnlicher Form werden Big-Data-Analytics-Lösungen bei der japanischen und britischen Bahn zur Optimierung der Verfügbarkeit, Pünktlichkeit und Auslastung der Züge eingesetzt.<sup>15,16</sup>

Mit Hilfe von Bewegungsdaten aus dem eigenen Mobilfunknetz sind Telekommunikationsanbieter in der Lage<sup>17</sup>, schnell und aktuell Reisezeitinformationen zu liefern. Damit lassen sich Verkehrsströme auf Autobahnen besser steuern und in der Folge die damit zusammenhängende Infrastruktur genauer planen. Die Swisscom hat eine »Mobility Insights Plattform« entwickelt, die das leistet.

### Umweltfreundliches Truckmanagement

DB Schenker hat eine Flottenlösung eingeführt, die zu signifikanten Reduktionen des Treibstoffverbrauchs und der CO<sub>2</sub>-Emissionen sowie gleichzeitig zu einem verbesserten Überblick für Disponenten führt.<sup>18</sup> Vernetzte Fahrzeugflotten erzeugen laufend Daten. Dabei entstehen schnell Datenvolumina in einem dreistelligen Terabyte-Bereich, in denen Optimierungspotenziale stecken. Die Flottenlösung sammelt kontinuierlich Fahrzeugbewegungsdaten und ermöglicht

<sup>14</sup> Vgl. (Lixenfeld, 2014).

<sup>15</sup> Vgl. (Dornacher, 2015).

<sup>16</sup> Im Juni 2015 hat die Deutsche Bahn ihre Digitalisierungsstrategie vorgestellt. Vgl. (Deutsche Bahn, 2015).

<sup>17</sup> Vgl. (Schlegel, 2015).

<sup>18</sup> Vgl. (Schwarzböck/Willenbrock, 2015).

dadurch die Berechnung der Fahreffizienz von Trucks.<sup>19</sup> Damit sind deutliche Senkungen des Kraftstoffverbrauchs und der daraus resultierenden Schadstoff-Emissionen möglich. Zusätzlich lassen sich durch die angefallenen Daten Korrelationen zwischen Fahrverhalten und notwendigen Reparaturen herstellen. Durch die ebenfalls effizientere Routenplanung ist auch eine genauere Prognose von Lieferzeiten möglich.

## Bessere Angebote für öffentlichen Nahverkehr

Vergleichbare Anwendungen entwickeln auch die Telekommunikations-<sup>20</sup> und andere Unternehmen<sup>21</sup>: Aktivitätsdaten aus dem Mobilfunknetz bieten die Möglichkeit, Verkehrsanalysen für unterschiedliche Verkehrsmittel mit hoher Detailtiefe durchzuführen. Für den öffentlichen Nahverkehr ergeben sich daraus neue Chancen, bessere Angebote zu schaffen. So lassen sich für den Verkehrsbetrieb mit Hilfe anonymisierter Daten aus dem Mobilfunknetz die Autofahrten zwischen unterschiedlichen Stadtteilen auswerten. Für vielbefahrene Strecken ergeben sich so immer bessere Alternativangebote. Eine Auswertung des Umsteigeverhaltens hilft zudem, Wartezeiten zu reduzieren.

Im Bereich Verkehr werden durch die Nutzung anonymisierter Daten aus dem Mobilfunk in naher Zukunft umfangreiche Möglichkeiten erschlossen: Bessere Daten für die Planung von Infrastrukturprojekten im Bereich Straße und Schiene, bessere Stauprognosen, eine bessere Taktung bei Zügen sowie Buslinien, die den ländlichen Raum besser erschließen – und vieles mehr.

## Verkehrs-Hubs – am Beispiel Smart Port Logistics im Hamburger Hafen

Mitten in der Stadt gelegen, nimmt der Hamburger Hafen rund ein Zehntel der gesamten Stadtfläche ein. Die Lage stellt die Hamburg Port Authority (HPA) vor besondere Herausforderungen. Sie bestehen darin, wachsende Warenströme auf diesem begrenzten Raum effektiv zu steuern, LKW-Standzeiten zu vermeiden und ein übergreifendes Lagebild der Ist-Situation für bessere Entscheidungen zu liefern.

Um die Verkehrsströme künftig miteinander zu vernetzen und effizienter zu gestalten, führte die HPA mit IT-Dienstleistern das Cloud-basierende Informations- und Kommunikationssystem

---

19 Per mobilem Gerät und GPS-Daten werden die jeweiligen Positionen eines Trucks sekundlich ermittelt. Auf einem zentralen Server werden diese Daten mit weiteren Informationen wie z. B. zum spezifischen Luftwiderstand des jeweiligen Lastwagens verknüpft. Die Auswertung dieser Kombination ergibt eine Aussage zur Effizienz des Fahrstils. Der Fahrer erhält die Information über ein Ampel-System auf seinem mobilen Endgerät: rot steht für eine ineffiziente, grün für eine effiziente Fahrweise.

20 Ein ausgereiftes Beispiel ist Motionlogic, eine Ausgründung aus den T-Labs der Deutschen Telekom: Vgl. (Weber, 2015).

21 Vgl. (Umbach/Kellner, 2015) sowie (Xu et al., 2015).

smartPORT logistics ein<sup>22</sup>, das den Andrang der täglich über 140000 Lkw-Fahrten im Hafen entzerrt und die Güterströme optimiert.<sup>23</sup>

Die Menge an Verkehrs- und Transportdaten ist riesig. Dabei gibt es viele verschiedene Beteiligte wie Hafenverwaltung, Speditionen, Transporteure, Parkflächen- und Terminal-Betreiber, die mit unterschiedlichen Systemen und Datenarten arbeiten.

Heute sind einzelne Akteure in sich optimiert, aber es gibt keine Optimierung des Gesamtsystems Ware/Verkehr. SmartPORT logistics ermöglicht es Häfen sowie Transport- und Logistikunternehmen, Transportaufträge in Echtzeit zu überwachen. Transportieren diese Organisationen Güter effizienter und sicherer, erhöhen sie auch die Zufriedenheit aller Beteiligten – insbesondere der privaten Verbraucher.

Im nächsten Schritt sollen Teilnehmer und Informationsquellen wie Informationen zur Verfügbarkeit der Container erweitert werden. Die HPA bietet daher einen Service-Marktplatz an, in dessen Datenangebote sich immer mehr Akteure der Hafenwirtschaft einklinken können.

### 4.3 Medizinische Forschung und Diagnostik

Big-Data-Anwendungen im Gesundheitswesen – zurzeit noch in den Anfängen – könnten Gesundheit und Wohlbefinden vieler Menschen bedeutend verbessern.<sup>24</sup>

Es entstehen derzeit Falldatenbanken<sup>25</sup> – z. B. ein Krebsregister an der Charité – für einige häufige sowie für schwere Erkrankungen, um bei neuen Patienten mit ähnlichen Symptomen die eine möglichst wirksame Therapie zu finden.

Die im Abschnitt 4.3 skizzierten Beispielanwendungen sind größtenteils bereits Teil unseres Lebens; einige befinden sich noch im Laborstadium. Sie benötigen personenbezogene, medizinische und dadurch besonders schützenswerte Informationen. Wenn es um Gesundheit und Leben geht, sollte der Staat die Rahmenbedingungen verbessern, damit die Patienten schnell von den Fortschritten in IT-gestützter Diagnostik und Therapie profitieren können (vgl. Leitlinie 3).<sup>26</sup>

22 Vgl. (Westermann/Müller, 2014) sowie (Bitkom, 2015, S. 55-56).

23 Verkehrssteuerung per Cloud generiert Big Data. 8.000 LKW bewegen sich pro Tag mit rund 140.000 Fahrten durch das Gelände. Über 100 Energieversorgungsunternehmen verkehren im Hafen. Der Hamburger Hafen ist der größte Eisenbahnhub Europas.

24 Für eine Gesamtsicht auf Big Data im Gesundheitswesen vgl. (Langkafel, 2015). Die Jahrestagung des Deutschen Ethikrates »Die Vermessung des Menschen – Big Data und Gesundheit«, Berlin, 21. Mai 2015, widmete sich dem Themenkreis Big Data und Ethik. Vgl. u. a. (Dabrock, 2014) und (Petri, 2015).

25 Oft handelt es sich um Big-Data-Lösungen.

26 Vgl. (Bohsem, 2015).

## Diagnose und Behandlung von Tumoren

Anfang Juli 2015 stellten das Potsdamer Hasso-Plattner-Institut und die Berliner Charité auf der Konferenz »Big Data in der Medizin« eine neue Lösung vor, mit der sich innerhalb kürzester Zeit die für einen Krebspatienten optimale Chemotherapie bestimmen lässt. Ohne Big-Data-Analysen würde das Wochen dauern. Die Analyse berücksichtigt die genetische Ausstattung des Krebspatienten und liefert Prognosen, wie Tumore auf bestimmte Medikamente ansprechen. So können auch Wirkstoffmengen genauer dosiert werden.

Auch am Nationalen Centrum für Tumorerkrankungen (NCT) in Heidelberg arbeiten Ärzte und Wissenschaftler an neuen Therapieformen. Das ist eine komplexe Aufgabe, denn jeder Tumor ist einzigartig. Das IT-Projekt »Medical Research Insights«<sup>27</sup> hilft den Forschern, individuelle Behandlungsmethoden für die Patienten zu entwickeln.

Die Erfolgsaussichten unterschiedlicher Behandlungsmethoden bei Tumorpatienten lassen sich oft nur vage vorhersagen – mit der Konsequenz, dass teure und nebenwirkungsbehaftete Therapien zu keiner Verbesserung des Gesundheitszustandes führen.

Komplexe Krankheiten benötigen komplexe Analysen. Es gibt über 100 verschiedene Formen von Krebs, wobei sich jede Form in ihrer Ausprägung und Entwicklung stark unterscheidet. Deshalb ist es umso wichtiger, dass betroffene Patienten eine schnelle, individuelle und auf ihren Fall zugeschnittene Behandlung erhalten. Dafür muss das NCT mit Bezug auf das Patientenprofil eine große Menge an Daten in Echtzeit auswerten (vgl. Leitlinie 3 und Leitlinie 8) – sowohl strukturierte<sup>28</sup>, als auch unstrukturierte<sup>29</sup> Daten. Alle Informationen fließen zentral zusammen. In kürzester Zeit lässt sich ermitteln, welche Therapie den größten Erfolg verspricht.

Big-Data-Technologien integrieren all diese Daten mit sehr großen Datenvolumen aus unterschiedlichen Quellen. Die beschleunigte Tumordiagnostik kann die Lebenserwartung der Tumorpatienten erhöhen.

## Hilfe für Infarkt-Betroffene

In der Vergangenheit haben Ärzte nur kurze Abschnitte von EKG-Daten analysiert. Nur selten waren Ärzte in der Lage, lange Ausdrücke zu studieren, die im Laufe eines Tages entstehen. Bei der Bewertung des Gesundungsprozesses von Infarkt-betroffenen kann es jedoch hilfreich sein, die gesamten Daten zu analysieren. Dabei haben US-amerikanische Wissenschaftler Hinweise in EKGs gefunden, welche Abweichungen von einem »normalen« Verlauf das Risiko vervielfachen,

<sup>27</sup> Basis ist SAP HANA, vgl. (Bitkom 2015, S.107f.).

<sup>28</sup> Tumordokumentationen, medizinische Akten, klinische Studien, ...

<sup>29</sup> Arztbriefe, MRT-Befunde, genetische Analysen oder Krebsregisterdaten, Behandlungshinweise, Versuchsstudien, Publikationen, ...

innerhalb eines Jahres einem erneuten Infarkt zu erliegen. So konnten neue Techniken für die Risikoanalyse entwickelt werden.<sup>30</sup>

## Personalisierte Medizin gegen Erbkrankheiten

Die Mukoviszidose ist eine vererbte Stoffwechselerkrankung – sie behindert die normale Lungenfunktion: Die Hälfte der Patienten wird keine 40 Jahre alt. Die Fortschritte bei der Gensequenzierung und bei Big-Data-Methoden erlauben es, personalisierte Arzneimittel herzustellen und betroffenen Menschen ein deutlich längeres und besseres Leben zu ermöglichen.

Die pharmazeutische Industrie will einzelnen Patienten gezielt helfen, indem sie klinische Daten in Verbindung mit genetischen Daten analysiert. Das Ziel dieser Analysen liegt darin, mehr über die Wirksamkeit und das Nebenwirkungsprofil von Medikamenten zu erfahren. Dadurch können die Ärzte jedem Patienten mitteilen, ob ein Präparat bei ihm wirkt und mit welchen Nebenwirkungen zu rechnen sein wird.

## Auswege aus der Makula-Degeneration

25 Millionen meist ältere Menschen auf der Welt leiden unter Makula-Degeneration, einer Gruppe von Augenerkrankungen, die in einem längeren Prozess zur Erblindung führt.

Die medizinische Forschung hat Wege identifiziert, wie man Betroffenen helfen kann: Die komplexen Signalvorgänge, die eine gesunde Netzhaut beim Betrachten eines Bildes im Gehirn auslöst, werden reproduziert, indem eine Spezialbrille Lichtsignale erzeugt. Diese Signale werden von Nervenzellen im Auge in elektrische Signale umgesetzt. Das Gehirn kann diese Signale verarbeiten – der Mensch kann wieder »sehen«.

Das komplexe Zusammenwirken einer ganzen Reihe von Hochtechnologien – Big Data bildet hier nur eine Gruppe – ist noch im Forschungsstadium, zeigt aber einen prinzipiellen Weg auf, Funktionen von Organen nachzubilden.

## Frühgeborenen-Überwachung

Kanadische Neonatologen haben festgestellt, dass es Fälle von Infektionen bei Frühgeborenen gibt, die durch medizinische Frühindikation ca. 24 Stunden vor dem Ausbruch der gefährlichen Infektion verhindert werden können. Durch die permanente Überwachung bestimmter Symptome und lebenswichtiger Parameter bei Frühgeborenen mit Hilfe modernster, medizinischer Sensorik lässt sich die Frühgeborenen-Sterblichkeit vermindern. Um die kritischen Warnsignale und Parameter-Änderungen des medizinischen Zustandes frühzeitig zu erkennen, müssen ununterbrochen tausende von Sensordaten pro Sekunde analysiert werden.<sup>31</sup>

<sup>30</sup> Vgl. (Smolan/Ewitt, 2013).

<sup>31</sup> Vgl. (Bitkom, 2012).

## Überwachung von Vitalparametern – Betreuung von Senioren und chronisch Kranken

Für die Erfassung und Verarbeitung detaillierter biometrischer Daten gibt es ein riesiges Spektrum von Anwendungen. Man denke an die Unterstützung von Sportlern beim Streben nach Höchstleistungen bis hin zur Betreuung einer alternden Bevölkerung.

Vielen älteren Menschen kann mit der Überwachung ihrer Vitalparameter geholfen werden, in ihrem häuslichen Umfeld zu bleiben – mit tiefgreifenden Folgen für Wohlbefinden und Lebensfreude sowie mit daraus resultierenden Einsparungen im Sozialwesen.<sup>32</sup>

Big Data und Cloud Computing liefern die Werkzeuge. Lösungen für solche Aufgabenstellungen sind längst Realität.<sup>33</sup> In einigen Ländern ist es bereits sehr populär geworden, solche Lösungen zur Gestaltung des eigenen Lebens einzusetzen, also z. B. die Balance zwischen Kalorienaufnahme und -verbrauch zu wahren.

Die meisten Patienten mit chronischen Erkrankungen wissen, dass sie in der Regel nicht mehr von ihrem Leiden befreit werden können. Patienten mit koronaren Erkrankungen (Herzinfarkt), Schlaganfall-Patienten oder Patienten, die unter Autoimmunerkrankungen leiden, geht es darum, sich im Verlauf ihres weiteren Lebens mit der Erkrankung zu arrangieren. Als Beispiel soll der Sport für Patienten mit koronarer Herzkrankheit dienen – neben der regelmäßigen Einnahme von Medikamenten eine lebensverlängernde Maßnahme. Viele Infarktpatienten geben jedoch die sportlichen Aktivitäten bereits kurz nach der erfolgreichen Rehabilitation wieder auf, da sie sich – alleine sporttreibend – ungeschützt fühlen und das Risiko einer Überanstrengung als zu hoch empfinden. Körpernahe Sensoren nehmen bei sportlich aktiven Patienten EKG-Signale und Vitaldaten ab. Die Daten werden über Mobiltelefone an einen Server übertragen, von Artefakten befreit, prädiktiv analysiert und mit Daten einer Vergleichskohorte verglichen. Bei der Detektion eines Zwischenfalls oder einer bedrohlichen Lage des Sportlers, die anhand der analysierten Daten festgestellt wurde, wird automatisch eine Ambulanz zum aktuellen Aufenthaltsort des Sportlers geschickt.

## Nutzen von Gesundheits-Apps

Seit 2007 hat sich der Smartphone-Markt durch nutzerfreundlichere Geräte und die stärkere Integration von Sensoren stark verändert. Dieser Trend hat sich noch durch Zusatzgeräte wie Fitnessarmbänder (Fitnesstracker) gesteigert, die leicht mit dem Smartphone zu koppeln sind.

---

32 In einer vom Bundesministerium für Wirtschaft und Energie geförderten Publikation wird auf dieses Geschäftsmodell verwiesen: »So kann [...] ein Betreiber von Diagnosegeräten Daten über den gesamten Bestand der in seiner Verantwortung betriebenen Geräte sammeln und auswerten und daraus neue Services, z. B. einen Diagnosevorschlag, generieren.« Vgl. (acatech, 2014, S. 21).

33 Vgl. (Bitkom 2013, S. 72f.).

Das Smartphone wird heute durch entsprechende Apps für Sportübungen, Lauftraining oder täglichen Gesundheitscheck zum persönlichen Kontrollzentrum und zur Steuerzentrale für den Einzelnen.

Im Detail können Zusatzgeräte für Pulsfrequenzmesser oder Pulsarmbänder mit Schrittzähler leicht per Bluetooth oder über den Kopfhöreranschluss gekoppelt werden. Selbst über die integrierte Kamera ist die Pulsmessung möglich. Indikatoren wie Aufwach- und Einschlafzeiten, zurückgelegte Schritte und gelaufene Kilometer lassen sich leicht erfassen und auswerten. Ableitungen wie generelle Fitness, verbrauchte Kalorien und Gesundheitszustand im Vergleich zu Referenzgruppen gleichen Alters, Geschlechts und Gewichts sind leicht umzusetzen.

Diese als Quantified Self bezeichnete Messung einzelner Vitalparameter des Menschen ist aber als persönliche Feedbacklösung nicht nur für den Einzelnen wertvoll, um Gesundheitszustand und Fitness zu überwachen.

Durch die digitale Übermittlung der anonymisierten Daten und Verarbeitung der einzelnen Detaildaten in Cloud-Lösungen entsteht ein Zusatznutzen. Die Verarbeitung tausender von Nutzungsprofilen und zehntausender Datenpunkte für einzelne Nutzer – Tageszeit, Wetter am Erfassungsort, Puls und Bewegungsmerkmale sowie Schlafverhalten des Anwenders – ermöglicht die maschinelle Mustererkennung als Big-Data-Anwendung. So können Signale für Erkrankungen und Früherkennung von gesundheitlichen Gefährdungssituationen genauer als je zuvor erfolgen.

Krankenkassen haben die Vorteile für ihre Mitglieder und die Kassen selbst erkannt und bieten bereits spezielle Tarife für Nutzer an, die sich entsprechende Apps der Krankenkasse auf ihrem Smartphone installieren.

Im Gegenzug zur Installation und dafür, dass die Krankenkassenmitglieder ihre Daten übermitteln, gibt es Einkaufs-Gutscheine und Rabatte. So vergibt zum Beispiel die Generali Versicherung für gesundes Verhalten Gutscheine an ihre Mitglieder. Der Anbieter Discovery leitet den Versicherten medizinische Tipps und Ratschläge über eine App weiter. Für die Steigerung der täglichen zurückgelegten Schritte gibt es Gutscheine. Einige Anbieter erwägen auch, Versicherungsbeiträge zu reduzieren (vgl. Leitlinie 6).

Durch Gesundheits-Apps und den Einsatz von Big Data zur zentralen Aggregation, Auswertung und Mustererkennung ergeben sich Vorteile für die Allgemeinheit, die Sozialeinrichtungen des Staates, für Unternehmen und auch für den Einzelnen (vgl. Leitlinie 12).



## 4.4 Öffentlicher Bereich

Im öffentlichen Bereich können Big-Data-Technologien bei der Erfüllung vieler öffentlicher Aufgaben helfen – wie etwa der Daseinsvorsorge, Stadtentwicklung und öffentliche Sicherheit.<sup>34</sup> Großes Potenzial bieten Wirtschaftsförderung und Smart-City-Ansätze.<sup>35</sup> Möglich sind auch Anwendungen zur Abschätzung von Gesetzesfolgen<sup>36</sup> oder zur Meinungsanalyse<sup>37</sup>.<sup>38</sup> Vor dem breiteren Einsatz von Big Data im öffentlichen Sektor<sup>39</sup> braucht es klare rechtliche und ethische Grenzen (vgl. Leitlinie 9).<sup>40</sup>

### Katastrophenmanagement

Die Hochwasserschutzzentrale der Stadt Köln setzt mit Partnern aus Wissenschaft und Wirtschaft das Projekt Smart-Data-Katastrophenmanagement (sd-kama) um.<sup>41</sup> Ziel des Projektes sd-kama ist es, eine Informationsplattform aufzubauen, die ein effizientes Katastrophenmanagement ermöglicht.<sup>42</sup> Einsatzleitung und Einsatzkräfte erhalten in Echtzeit – im Krisenfall zählt jede Minute – einen Überblick über die aktuelle Hochwasserlage (betroffene Gebiete, Ausmaß der Schäden, Position der Rettungskräfte) sowie eine Prognose über die weitere Entwicklung (Katastrophenverlauf, Wettervorhersage). Diese Informationen erlauben es, lokale Gegebenheiten zu berücksichtigen, Plausibilitäten zu prüfen und Entscheidungen zu priorisieren. Momentan ermöglichen nur teure und zeitaufwändige Hubschrauberflüge ein solch zusammenhängendes Lagebild.

Das Herzstück von sd-kama bildet eine Smart-Data-Plattform, deren technischen und sicheren Betrieb das Smart Data Innovation Lab übernimmt. Auf der Plattform werden Video- und Bilddaten (YouTube, Instagram), Fernerkundungsdaten (Satelliten), Sensordaten (Pegelstände) sowie Bestandsdaten (Kataster) zusammengeführt und analysiert. Die Informationen werden räumlich visualisiert und in das Flutinformations- und Warnsystem der Hochwasserschutzzentrale der Stadt Köln eingebunden.

---

34 Bei Organisationen wie dem Deutschen Patent- und Markenamt gehört der Big-Data-Einsatz zum Kerngeschäft – vgl. (Rülke, 2013).

35 Vgl. (Eckert et al, 2014).

36 Vgl. (Bitkom 2015a), insbesondere Abschnitt 4.2 über »Kognitive Systeme zur Verbesserung öffentlicher Entscheidungen? Einsatz in der Gesetzesfolgenabschätzung«.

37 Vgl. (Nold, 2013).

38 Vgl. (Eckert et al, 2014).

39 Vgl. dazu Kapitel 5.

40 Vgl. Fußnote 35.

41 Vgl. (Willkomm, 2014).

42 Vgl. auch Abschnitt 4.5.

## Cybersecurity

Einen wichtigen Beitrag zur Sicherheit öffentlicher IT-Infrastrukturen leistet die Überwachung der im IT-Betrieb generierten und mit Big-Data-Ansätzen ausgewerteten Maschinendaten. Dazu gehört das Sammeln, die zentrale Aggregation, die Langzeitspeicherung, die Analyse der Logdaten sowie die Suche in den Logdaten und daraus generierte Berichte.

Die Quellen und Formate der Logdaten sind vielfältig. Die Zentralisierung macht Daten verfügbar und im Zusammenhang auswertbar und bildet daher die Grundlage für die Überwachung und Kapazitätsplanung des IT-Betriebs. Das gilt genauso für die Durchführung von Sicherheitsmaßnahmen oder die Protokollierung der Einhaltung von Richtlinien und Vorschriften.

Verdächtige Veränderungen oder Vorgänge möglichst rasch – am besten in Echtzeit – zu erkennen, ist der Anspruch der IT-Sicherheit. Die Auswertung von Logdaten bietet hier einen sehr guten Ansatzpunkt. Unverzichtbar wiederum ist sie für den Nachweis oder die Spurensuche nach Sicherheitsvorfällen.

Die Korrelation von Logdaten aus verschiedenen Quellen holt Zusammenhänge ans Licht, die ansonsten verborgen sind. Big-Data-Technologien lassen sich in vielen Bereichen wie etwa für das Aufspüren von hochentwickelten Bedrohungen, Insider-Angriffen oder der missbräuchlichen Verwendung von Zugangsdaten nutzen.

## Kriminalitätsvorbeugung

Innenministerien einzelner Bundesländer prüfen die Möglichkeiten, mit Methoden des Predictive Policing bestimmten Delikten wie z. B. Wohnungseinbrüchen oder dem Entwenden hochwertiger Pkw vorzubeugen.<sup>43</sup> Die USA sowie einige europäische Länder (Niederlande, Schweiz) sind Vorreiter in diesem Bereich.

Die Polizei prognostiziert mit Big-Data-Analysen Tatorte und Delikte. Auf Basis gesammelter historischer Daten entwickelt sie dann Gegenstrategien. Dazu gehören beispielsweise die verstärkte Präsenz von Ordnungshütern an Brennpunkten des kriminellen Geschehens.

Predictive Policing unterstützt die Polizei im Streifendienst.

## IT-Sicherheit in künftigen Energienetzen

Big-Data-Ansätze tragen zur Gewährleistung der Sicherheit in künftigen Energienetzen bei.<sup>44</sup>

<sup>43</sup> Vgl. Dieter Schürmann, Landeskriminaldirektor im Ministerium für Inneres und Kommunales des Landes Nordrhein-Westfalen, auf einer Podiumsdiskussion über Cognitive Computing am 25. November 2014, Bitkom Trendkongress.

<sup>44</sup> Vgl. (Kemeröz/ Kremer 2015) sowie Abschnitt 4.1.

Intelligente Messsysteme (Smart Meter und Smart Meter Gateways) dienen dazu, den Energieverbrauch in jedem Haushalt in kurzen Zeitabständen zu messen, anzuzeigen und dem Energieversorger zu übermitteln. Auf Basis dieser Daten steuern Energieversorger gezielt und individuell Erzeugeranlagen und Verbraucher.

Um zeitnah auf Schwankungen im Stromnetz reagieren zu können, müssen hunderttausende Smart Meter Gateways im Stand-by-Modus auf verschlüsselte Befehle warten und diese dann unmittelbar an intelligente Steuergeräte weitergeben, die Erzeuger- oder Verbraucheranlagen an bzw. abschalten. Erst damit wird ein schneller und direkter Datenaustausch ohne die vorherige Verbindungsaufbauprozedur möglich. Bis zu zwölf Millionen Smart Meter Gateways sollen bis 2022 verbaut werden und auf Basis der vorgeschriebenen sicheren Infrastruktur kommunizieren.<sup>45</sup>

## Umwelt-Monitoring

Die Biodiversität stellt eine der wichtigsten Grundlagen unseres Lebens dar. Das Monitoring der Biodiversität hat als Zielsetzung, die Auswirkungen von Klimawandel und Landnutzungsänderungen auf die Vielfalt der Arten zu verstehen. Für ein aussagefähiges Monitoring sind ausreichende Daten und langfristige Beobachtungen notwendig. Für das Ökosystem der tropischen Regenwälder betreibt das TEAM-Netzwerk (Tropenökologie, Bewertung und Monitoring) das weltweit größte System von Kamerafallen (14 Länder, ca. 1.000 Kamerafallen auf einer Fläche von über 2.000 km<sup>2</sup>). Diese Kamerafallen erfassen systematisch das Vorkommen von terrestrischen tropischen Säugetieren und Vögeln. Die primären Kameradaten liefern Trends zum Vorkommen der erfassten Tierarten. Zusammen mit weiteren Einflussfaktoren wie Klima und Landnutzung wird mithilfe statistischer Verfahren der Wildlife Picture Index ermittelt. Dank dieses Monitorings können staatliche Stellen leichter ihren Überwachungs- und Schutzpflichten nachkommen.<sup>46</sup>

## 4.5 Vorhersage von Krisensituationen

Empirische Forschung in den Natur- oder Sozialwissenschaften ist auf die Verfügbarkeit von Daten zur Theoriebildung angewiesen. Der Einsatz von Big-Data-Technologien ermöglicht genauere Theorien und damit die Beantwortung von Fragestellungen mit bisher nicht bekanntem Detailgrad, insbesondere auch präzisere Vorhersagen von Krisensituationen. Beispiele findet man in der Klimaforschung bei der Vorhersage von Extremwetterlagen sowie den Wirtschaftswissenschaften bei der Vorhersage von Finanzkrisen.

Unter dem Oberbegriff »komplexe Systeme« entwickelt sich aktuell in den Wissenschaften ein neues Paradigma, das in Ergänzung bekannter Differentialgleichungsmodelle neuartige mathematische Modelle wie Netzwerke oder agentenbasierende Systeme zur Beschreibung von Vorgängen in Natur und Gesellschaft einsetzt. Parametrisiert durch große Datenmengen, erlauben

<sup>45</sup> Vgl. (Bitkom, 2015).

<sup>46</sup> Vgl. (Bögershausen, 2014).

diese Modelle Vorhersagen, insbesondere von Extremsituationen mit Auswirkungen für die gesamte Gesellschaft. Beispiele sind das Eintreten von Naturkatastrophen – Überflutungen<sup>47</sup>, Trockenzeiten, Erdbeben – oder wirtschaftlichen Extremsituationen (Finanzkrisen). Durch die zunehmende Verfügbarkeit von Daten, die Natur und Gesellschaft in bisher unbekannter Auflösung beschreiben, ergibt sich die Möglichkeit präziserer Theorien und damit besserer Vorhersagen. Diese ermöglichen es Politik, Wirtschaft und Gesellschaft, früher auf zu erwartende Extremsituationen zu reagieren und dadurch gewaltige Schäden abzuwenden.

## Klimaforschung

In der Klimaforschung ist der Umgang mit großen Datenmengen bereits seit Jahrzehnten üblich. Datenbanken, die das globale Klima und lokale Wetterhistorie erfassen, sind notwendig, um Vorhersagemodelle zu parametrisieren und ihre räumliche und zeitliche Auflösung zu erhöhen. Nur so lassen sich die Interaktionen zwischen globalen und lokalen Dynamiken erfassen und für frühzeitige Warnungen nutzbar machen. Beschränkungen in der Speicher- und Rechenkapazität limitieren hier nach wie vor die Genauigkeit. Durch den Einsatz von Big-Data-Technologien in Kombination mit neuartigen Netzwerkmodellen ist es beispielsweise Wissenschaftlern am Potsdamer Institut für Klimafolgenforschung gelungen, 90 Prozent der monsunbedingten extremen Wetterereignisse mit teilweise verheerenden Auswirkungen in Südamerika wie Fluten oder Erdbeben vorherzusagen.<sup>48</sup>

Ähnliche Fortschritte ermöglicht Big Data bereits jetzt bei der Abschätzung der Effekte von klimawandelbedingten Extremwittersituationen auf gefährdete Infrastrukturen wie Straßen, Strom- oder Kommunikationsnetzwerke sowie unter Umständen bald bei der Früherkennung von Erdbeben und Tsunamis.<sup>49</sup>

## Vorhersagen von Finanzkrisen

Mit ähnlichem Grundansatz verfolgen ambitionierte sozial- und insbesondere wirtschaftswissenschaftliche Forschungsprojekte das Ziel, Muster beim Auftreten von Finanzkrisen zu identifizieren und dadurch frühzeitig Warnungen zu geben. Die 2008 einsetzende Wirtschafts- und Finanzkrise hat gezeigt, dass die Finanzmärkte zu Instabilität neigen und in kurzer Zeit extreme Fluktuationen entstehen können, welche die Weltwirtschaft enorm beeinträchtigen. Obwohl es gewichtige Einwände gibt, die gegen eine präzise Vorhersagbarkeit der Dynamik von Finanzmärkten sprechen, ist es im Interesse der Allgemeinheit, zukünftig solche Extremereignisse zu vermeiden. Unabhängig von der Frage, ob dies durch stärkere Regulation gelöst werden kann, ist klar, dass dazu ein besseres Verständnis der zugrunde liegenden Dynamiken notwendig ist. Auch hier ist Big Data relevant und wird bereits in der Forschung erprobt: Forscher der ETH Zürich haben

47 Vgl. das im Abschnitt 4.4 vorgestellte Projekt der Stadtverwaltung Köln zum Hochwasserschutz.

48 Vgl. (PIK, 2014).

49 Vgl. <http://www.quakehunters.com>, Satellite Big Data, Earth Observation & Internet System to reduce the risk and damage of earthquakes.

beispielsweise in der FuturICT-Initiative<sup>50</sup> die Konstruktion einer weltweiten Plattform für die Erfassung und das Verständnis der digitalen Spuren menschlichen Verhaltens vorgeschlagen. Dieses Observatorium soll Wissenschaftler in die Lage versetzen, nicht nur Finanzkrisen, sondern auch politische Unruhen wie den arabischen Frühling und dadurch bedingte soziale Extremsituationen vorherzusagen. Während dieses EU-weit ausgelegte wissenschaftliche Großprojekt bislang keine Förderung erhalten hat, existieren bereits einzelne Initiativen wie das Financial Crisis Observatory<sup>51</sup>, die sich dem Finanzmarkt widmen.

## 4.6 Bildung und Qualifizierung

Intelligente Bildungsnetze, ausgestattet mit digitalen Inhalten, gebrauchstauglichen Lernumgebungen und Möglichkeiten zur sozialen Vernetzung und Kooperation über Institutionengrenzen hinweg, unterstützen heute sowohl den Prozess des lebenslangen Lernens als auch das personalisierte Lernen des Einzelnen und von Gruppen in der formalen und informellen Bildung und Qualifizierung. Die Erfassung, Speicherung und Verarbeitung großer Datenmengen über das individuelle Lernverhalten und die Veränderung personaler Kompetenzen in hybriden Welten hat dabei eine lange Tradition, ist verbindendes Element und wird für Erklärung und Prädiktion genutzt. Learning Analytics-Methoden kommen zum Einsatz, welche einerseits die computerbasierte, automatisierte Anpassung digitaler Bildungsinhalte, Lernmethoden und Bildungsumgebungen an den Lernenden unterstützt, andererseits eine marktgerechte, nachfrageinduzierte Realisierung von Unterrichts-, Studien- und Weiterbildungsangeboten für das lebenslange Lernen ermöglicht.

### Personalisierte Lernumgebungen

Zentrales Element personalisierter Lernumgebungen in einer digitalen Welt ist der Lernende selbst. Ob formale Bildungsarrangements in Schule oder Hochschule oder informelle Qualifizierungssettings in der Lern- und Forschungsfabrik, am Arbeitsplatz im Unternehmen oder zu Hause – im Mittelpunkt steht der aktive, selbstgesteuerte Lernende mit individuellen Voraussetzungen, eigenen Lernpräferenzen, Lernwegen, Lernzielen und letztlich auch individuell präferierten medialen Lernangeboten.

Intelligente Bildungsnetze unterstützen den Lernenden in personalisierten Lernumgebungen durch die datenbasierte, computergenerierte Gestaltung und Anpassung von Zielstellung, Inhalt und Methodik. Dabei werden die Auswahl und Präsentation digitaler Lernobjekte und das didaktisch-methodisch begründete Vorgehen im Bildungsprozess von bekannten, lernrelevanten Merkmalen des Lernenden beeinflusst und auf diese Weise der Lernprozess hochgradig individualisiert. Grundlage jedweder Form der Gestaltung und Anpassung sind dabei alle relevanten Daten des Lernenden, der Lernobjekte, der Lernumgebung, Veränderungen über die Zeit, die

<sup>50</sup> Vgl. [www.futurict.blogspot.de](http://www.futurict.blogspot.de)

<sup>51</sup> <http://www.er.ethz.ch/financial-crisis-observatory/description.html>

letztlich unterschiedliche Formen der Adaptivität ermöglichen. Schnittstellen-Adaptivität (Modifikationen der Benutzeroberfläche), statistische Adaptivität (Modifikation der Lernumgebung an Wissensstand, Lerngewohnheit, Inhaltspräferenzen) und dynamische Adaptivität (Modifikationen aufgrund des Nutzungs- und Lernpfades in einer Lernumgebung) sind konkrete Ausprägungsformen.

Intelligente Wissensdienste, die in zunehmendem Maße, nicht zuletzt durch die Entwicklungen rund um das Internet der Dinge, Dienst und Daten sowie Industrie 4.0 induziert, in Unternehmen der Wirtschaft und Industrie, aber auch schon seit Ende der 1990er Jahre in Schulen und Hochschulen eingesetzt werden, können beispielhaft für die Adaptivität und Personalisierung von Lernumgebungen an den Einzelnen betrachtet werden. Durch Analyse und Nutzung großer Datenbestände aus darin integrierten Modellbetrachtungen von Wissen, Lerner und Bildung werden Anpassungen auf Basis von Methoden der Künstlichen Intelligenz ermöglicht. Das Wissensmodell besteht aus dem in Lernobjekte strukturierten und mit Metadaten annotierten Lehr-Lern-Material. Das Lernermodell wird zur datenbasierten, individuellen Fortschrittskontrolle herangezogen. Und das Bildungsmodell erzeugt die Passung zwischen dem Wissensmodell und Lernermodell basierend auf Daten u.a. zu Zielen, Präferenzen, Erfahrungen. Exemplarisch können die Projekte APPSist<sup>52</sup> und ALINA<sup>53</sup> angeführt werden.

## Educational Data Mining

Ein besseres Verständnis über das Verhalten von Lernenden in heterogenen Bildungs- und Qualifizierungsumgebungen sowie über den Lebenslauf hinweg steht seit Mitte des letzten Jahrzehnts im Fokus des Educational Data Mining. Hier werden Methoden entwickelt, erprobt und validiert, die einzigartige, spezifische Daten in großem Ausmaß von digitalen Bildungsszenarien nutzen. Daten können dabei von Lernumgebungen, kollaborativen Werkzeugen, digitalen Lernobjekten oder auch vom Lernenden selbst kommen, etwa seine Kompetenzen und seine Expertise, sein Lernfortschritt, aber auch sein Verhalten in Sozialen Netzwerken oder beim Besuch von Informations- und Wissensplattformen, die frei im Internet zur Verfügung stehen.

Seit 2012 werden in Deutschland verstärkt die Interpretation von Lernenden im formalen Hochschulkontext betrachtet und unter dem Begriff »Learning Analytics« zusammengefasst. Lernfortschritte in digitalen Bildungs- und Qualifizierungsumgebungen werden gemessen, Studierendendaten erhoben, zukünftige Leistungen vorausgerechnet und potenzielle Problembereiche in der Bildung aufgedeckt. Dabei werden Daten aus Campus-Management-Systemen, Learning-Management-Systemen, Online-Assessment-Verfahren ebenso genutzt wie der individuelle Nutzen von Informations- und Wissensdatenbanken und -systemen im privaten Kontext oder in Sozialen Netzwerken. Daten werden aufeinander bezogen, aggregiert und analysiert. Nicht selten stehen neben pädagogischen Überlegungen auch ökonomische und politische Aspekte im Hintergrund. So geht es auch um die Effizienz- und Effektivitätsbetrachtung von Maßnahmen,

52 Vgl. <http://www.appsist.de> (abgerufen am 31. August 2015).

53 <http://www.qualifizierungdigital.de/projekte/laufende-projekte/alina.html> (abgerufen am 31. August 2015).

um Wirtschaftlichkeitsanalysen mit Bezügen zu Business Intelligence und Möglichkeiten des Bildungscontrollings für Bildungsinstitutionen.

## 4.7 Finanzdienstleistungen

Bei Finanzdienstleistern werden Big-Data-Technologien zurzeit in erster Linie in der Kreditvergabe sowie bei der Eindämmung von Betrug<sup>54</sup> und Geldwäsche eingesetzt – für Verbraucher sehr wichtige Einsatzgebiete.<sup>55</sup>

### Einsatz von Big Data bei der Kreditbeschaffung

Im traditionellen Bankgeschäft werden seit vielen Jahren analytische Verfahren eingesetzt. Diese erleichtern und objektivieren Kreditentscheidungen etwa beim Hauskauf oder im Bereich der Verbraucherkredite. Big-Data-Analysen erlauben nun wesentlich genauere und schnellere Entscheidungen. Durch die Auswertung von typischem Verhalten von Kreditausfällen lassen sich gezieltere Vorhersagen zur Kreditwürdigkeit von Kunden treffen. Die damit gebauten Modelle verhindern auch eine allzu schnelle Ablehnung aufgrund des Fehlens von starren Voraussetzungen, sondern geben auf einer viel breiteren Basis ein genauere Bild über das voraussichtlich Zahlungsverhalten und damit des Risikos, das die Bank eingeht.

Eine Reihe von Start-up-Unternehmen in der Branche nutzen diese Big-Data-Technologien und Datentöpfe aus, um neue Zielgruppen zu erschließen. Selbstlernende Algorithmen, die Informationen etwa aus Facebook oder anderen Sozialen Netzwerken mit Einwilligung des Kunden einholen, erlauben eine automatisierte und schnelle Entscheidung über Kredite. Häufig handelt es sich dabei um Volumina, die im traditionellen Bankgeschäft zu gering und mit zu hohen Kosten für die Banken verbunden sind. Daher profitieren insbesondere bisher nicht mit Bankdienstleistungen versorgte kleine und mittelständische Unternehmen sowie Verbraucher in Schwellen- und Entwicklungsländern von dieser Entwicklung.

Die Verbesserung des Kreditprozessen sorgt insgesamt dafür, den reibungslosen Ablauf der Wirtschaft zu stärken sowie Verbraucherrechte zu wahren.

---

54 Auf dem 2. Bitkom Big Data Summit wurde von einer am Fraunhofer SIT entwickelte Methode zur Datenanalyse in einem viele Millionen Einträge umfassenden Praxisfall bei Ernst & Young Fraud Investigation berichtet, mit der bilanzbeeinflussende Unstimmigkeiten in den Finanzdaten einer großen Versicherungsgesellschaft aufgedeckt wurden – vgl. (Winter/Mollat, 2014).

55 Big-Data-Konzepte unterstützen auch das Risikomanagement bei Finanzdienstleistern. Risikohinweise für den Finanzmarkt sind in den täglichen hunderttausenden Nachrichten, Chats und Messages enthalten, die sich inhaltlich analysieren und bewerten lassen. Vgl. (Klemm/Eckl 2015).

## Aufdeckung von Versicherungsbetrug

Versicherungsbetrug und Geldwäsche locken raffiniert vorgehende kriminelle Kreise an, die sich häufig sehr wirksam der Aufdeckung ihrer Betrugereien entziehen. Daraus erwächst die Herausforderung, durch Netzwerk-Analyse Zusammenhänge bei Finanztransaktionen in Echtzeit zu erkennen.<sup>56,57</sup> Hier kommen Graphdatenbanken<sup>58</sup> zum Einsatz. In Kombination mit den bewährten Betrugsbekämpfungsmethoden bieten sie Unternehmen umfassenderen Schutz vor Missbrauch. Auf diese Weise reduzieren Banken und Versicherungen den Kapitalausfall durch Missbrauchsschäden und können so ihre Einnahmen verbessern bzw. die Preise für ihre Dienstleistungen senken und damit ihren Kunden attraktivere Angebote unterbreiten.

## 4.8 Landwirtschaft

Die Bereitstellung von Nahrungsmitteln für eine auch weiter rasant steigende Weltbevölkerung ist eine eminent politische Aufgabe. Insofern sind die Möglichkeiten des Big-Data-Einsatzes in der Landwirtschaft von enormer gesellschaftlicher Relevanz, was ausgewählte Beispiele zeigen.

### Steuerung von Landmaschinen

Als fortgeschrittener Nutzer von Big-Data-Technologien hat sich die Claas KGaA mbH einen Namen gemacht. Das Unternehmen stattet seine Landmaschinen mit Sensoren und Messsystemen aus, die einen Datenstrom erzeugen. Die Daten aus dem Entwicklungs- und Produktionsprozess werden durch Telematics- und Betriebsdaten sowie Daten interner Systeme für die Ersatzteilbestellung oder Schadensmeldungsbearbeitung ergänzt. Aus der Erweiterung des zentralen Datenpools durch Zusatzinformationen wie z. B. Wetterhistorien oder Bodenbeschaffenheitsdaten entsteht ein Data Lake mit umfassenden Big-Data-Einsatzszenarien<sup>59</sup>.

Zur Ausweitung der Big-Data-Kompetenz im Unternehmen arbeitet die Claas KGaA mbH am BMBF-Förderprojekt »Agata – Analyse großer Datenmengen in Verarbeitungsprozessen« mit, das das Fraunhofer-Anwendungszentrum Industrial Automation (IOSB-INA) koordiniert. Es zielt auf die Auswertung großer Datenmengen, die in Prozessen der Industrie und Landwirtschaft anfallen. Im Projekt wird ein selbstlernendes Assistenzsystem entwickelt, das durch die Beobachtung komplexer Verarbeitungsprozesse in Industrie und Landwirtschaft Zusammenhänge ermittelt und so Fehler, Anomalien und Optimierungsbedarf automatisch erkennt.<sup>60,61</sup>

56 Vgl. (Bitkom, 2015).

57 Vgl. (Stolberg, 2015).

58 Graphdatenbanken sind zur Abfrage verschlungener Netzwerke konzipiert – in Echtzeit, während der Betrug stattfindet, und ohne vorher zu wissen oder zu ahnen, wer mit wem verbunden sein könnte.

59 Vgl. (Porzig/Vospeter, 2015).

60 Vgl. (Engeser, 2014).

61 Vgl. (Schreier/Litzel, 2014).



Ziel ist es, mit Hilfe der Big-Data-Analysen einen höheren landwirtschaftlichen Ertrag bei einer Verringerung des Ressourceneinsatzes zu erreichen. Das wird in der Praxis etwa dadurch erreicht, dass Düngemittel gezielter eingesetzt und Bewässerung besser auf die Wetterverhältnisse abgestimmt ist

### Optimierung von Wasser- und Düngerzufuhr für landwirtschaftliche Kulturen – am Beispiel der Weinerzeugung in Kalifornien

Das kalifornische Weingut Shafer<sup>62</sup> im Napa Valley zählt zu den Pionieren bei der datengestützten Industrialisierung des Weinbaus.<sup>63</sup> So nutzt das Unternehmen, das seit 1972 Wein anbaut und weltweit vertreibt, seit einigen Jahren unterschiedliche digitale Technologien, um die Bewässerung und Düngemittelzufuhr auf seinen Anbauflächen zu optimieren. Denn so lassen sich bei gleichbleibender Qualität deutlich höhere Erträge pro Hektar erwirtschaften. In den Weinbergen von Shafer kommen Wetterstationen zum Einsatz, die das Wetter in den Weinbergen per Sensor erfassen, über WLAN oder Satellit an einen zentralen Server senden und dann per Web-Schnittstelle für den Winzer bereitstellen. Die Wetterstationen messen unter anderem Temperatur, Luftfeuchtigkeit, Sonneneinstrahlung und Bodenfeuchtigkeit. Zudem bietet das System die Auswertung historischer Wetterdaten, anhand derer der Winzer langfristige Trends erkennen und seine eigenen Entscheidungen justieren kann. Zudem kommt im Weinberg von Shafer eine Sensor-Technologie zum Einsatz, die direkt an den Weinreben Daten aufnimmt und per Wifi in ein zentrales Analysesystem leitet. Anhand des Beispiels wird deutlich, dass auch in einer der ältesten Branchen der Welt die Digitalisierung nicht mehr aufzuhalten ist und große Datenmengen anfallen, die althergebrachte Traditionen in Frage stellen oder ergänzen. Die kalifornischen Winzer der nächsten Generation setzen zunehmend auf eine Verbindung von Erfahrung und Data Science, um gute Weine zu produzieren und hohe Erträge zu erwirtschaften.

## 4.9 Connected Car – aus Sicht privater Nutzer

Sensor-, Cloud- und Big-Data-Technologien bilden die Grundlage für autonome Fahrzeuge und lassen tiefgreifende Veränderungen in der Ausgestaltung von Mobilitätslösungen erwarten.

Zurzeit testen führende Automobilhersteller Modelle, die in fünf oder sechs Jahren auf den Markt kommen werden. Die futuristischen Fahrzeuge lenken, fahren und bremsen allein, gesteuert von Kameras, Sensoren und Laserscannern. Damit hat die Digitalisierung endgültig die Automobilindustrie erobert. Big-Data-Technologien ermöglichen die schnelle Verarbeitung und Auswertung vielfältiger Informationen sowie den Einsatz riesiger Datenmengen für die Prognose des Verkehrsflusses (vgl. Leitlinie 7). Das revolutioniert die Verkehrssteuerung. Das Auto wandelt sich stärker zum Arbeits- oder Lebensraum.

62 <http://www.shafervineyards.com/story/index.php> (abgerufen am 31. August 2015).

63 Vgl. <http://www.theverge.com/2013/8/19/4636190/winemakers-turn-to-web-sensors-as-some-still-use-feetp> (abgerufen am 31. August 2015).

Neben dem Austausch von Informationen zwischen zentralen Stellen und den Fahrzeugen (vgl. Leitlinie 8) können diese auch direkt miteinander kommunizieren (Car-to-Car-Communication) und so die Vorfahrt regeln oder vor Hindernissen warnen. Autonome Autos haben somit weitreichende Folgen auf die Gesellschaft (vgl. Tabelle 1):

**Tabelle 1: Gesellschaftliche Implikationen autonomer Fahrzeuge**

Folgen	Erläuterung
<b>Mehr Sicherheit</b>	Auch Maschinen sind mit technischen Mängeln behaftet, aber insgesamt ist die Fehlerquote der Maschinen deutlich geringer als die der Menschen. Schon die heutigen Assistenzsysteme haben den Verkehr deutlich sicherer gemacht. Umgekehrt ist es allerdings sicher so, dass bei Unfällen der Autopilot kritisch hinterfragt und untersucht wird, ob z. B. der Automobilhersteller ein Mitverschulden hat.
<b>Weniger Staus</b>	Wenn nur zehn Prozent aller Fahrzeuge regelmäßig ihre Geschwindigkeit und Position weitergeben, können autonome Autos vorausschauend fahren und Staus vermindern. Die Kapazitäten der Fahrbahnen werden deutlich steigen.
<b>Weniger Verbrauch</b>	Indem autonome Autos optimal schalten, Gas geben und bremsen, können sie mit deutlich weniger Treibstoff auskommen.
<b>Mehr Komfort</b>	Autonome Autos können die Nachteile von Carsharing gegenüber der Nutzung des eigenen Autos eliminieren. Wenn sich Mietwagen per Smartphone bestellen lassen und sich am Ziel selbst eine Parklücke suchen, spart der Benutzer sowohl den Weg zur Carsharing-Station als auch die Suche nach einem Abstellplatz.

## 4.10 Digitale Einkaufswelten

Das Internet der Dinge und Big Data dringen schrittweise von der Produktionskette bis an die Ladentheke. Big Data gibt dem (Online-)Händler Werkzeuge in die Hand, die potenziellen Käufer mit individuellen Angeboten zu adressieren und bis zur Kaufentscheidung zu begleiten. Neueste Entwicklungen tragen zur Verminderung von Retouren bei<sup>64</sup> – im Interesse des Einzelnen und der Gesellschaft gleichermaßen.

### Outdoor Analytics

Big Data erleichtert die Mobilitätsforschung und -kontrolle.<sup>65</sup> Wo Personenströme von Bedeutung sind, wird verlässliche Planung zur Herausforderung:

- Wie viele Menschen gehen pro Tag an meinem Schaufenster vorbei?
- Wo ist der (Nah-)Verkehr besonders ausgelastet?
- Wo ist der beste Platz für eine Werbetafel?

<sup>64</sup> Vgl. (Bruysten, 2015).

<sup>65</sup> Erinnert sei an die Massenpanik bei der Love-Parade in Duisburg im Jahre 2010 mit 21 Toten.

Solche Fragen ließen sich bisher nur mit aufwendigen Handzählungen und rückblickend beantworten. Mit Big Data in Verbindung mit Lokationsdaten (vgl. Leitlinie 10) sind schnelle Antworten möglich. Informationen über Verkehrsströme – angereichert mit zusätzlichen Informationen – lassen sich für die zielgerichtete Verbraucheransprache nutzen. Den Verbrauchern erleichtern solche Informationen die Orientierung in für ihn neuen Regionen. Entsprechend ihren Vorlieben und aktuellen Bedürfnissen erhalten sie Informationen über Hotels, Restaurants, aktuelle Angebote oder Freizeitmöglichkeiten. Früher oftmals zeitaufwendige organisatorische Fragen lassen sich nun zügig erledigen und neue kurzfristige Angebote in Erfahrung bringen.

### Indoor Analytics – Optimierte Einkaufswege

Einkaufszentren werden mitunter als Irrgärten empfunden. Die Orientierung per Smartphone funktioniert bislang nicht. Neue Techniken können nun Abhilfe verschaffen. Die Verbraucher vermeiden Umwege, wenn sie in Kauf nehmen, dass die Anbieter Bewegungsprofile erstellen.

Unternehmen mit Verkaufsräumen stellen sich immer wieder die gleichen Fragen:

- Welchen Weg nehmen die Verbraucher durch das Geschäft?
- Vor welchem Aufsteller bleiben sie wie lange stehen?
- Welches Angebot zieht die meisten Interessenten an?

Big Data in Verbindung mit Lokationsdaten ermöglicht die Auswertung von Kundenbewegungen in Verkaufsräumen oder Einkaufszentren in Echtzeit und somit die Optimierung von Verkaufsräumen und Sortimenten. Der Verbraucher profitiert von Informationen, die ihm – entsprechend seinen Vorlieben oder angegebenen Wünschen – zur Verfügung gestellt werden: Wo und zu welchen Konditionen sind bestimmte Produkte verfügbar. Ein Routenplaner sichert kurze Wege.

## 4.11 Online-Marketing und Nutzen von Big Data für Verbraucher

Big Data trägt zur Erhaltung eines kostenfreien, da werbefinanzierten Angebots-Ökosystems im Internet bei. Verbraucher profitieren von qualitativ hochwertigen und kostenfreien Inhalten, Dienstleister von optimierten Erträgen aus der Bereitstellung ihrer Services. Die Erfassung, Aggregation und Aussteuerung von Daten und Werbeinhalten erfolgt datenschutzkonform und anonymisiert.

Seit Verbraucher das Internet stärker nutzen, steht die Frage im Raum, wie die Bereitstellung von Informationen und Services wirtschaftlich genutzt werden kann. So sollte für Anbieter ein Anreiz bestehen, qualitativ hochwertige und aktuelle Informationen sowie wertvolle und nutzenstiftende Services bereitzustellen.

Online-Marketing ist ein substantieller Bestandteil des Internets, da die meisten Nutzer kostenfreie Inhalte wünschen. Als Gegenwert stehen Werbeeinblendungen in unterschiedlicher Form. Ohne Werbung ist also kein kostenfreier Inhalt im Internet denkbar (vgl. Leitlinie 11).

Viele Jahre war Online-Werbung als Baustein für die Finanzierung von Online-Diensten kein brauchbares Instrument, weil die technischen Lösungen für die zielgerichtete Ansprache der Nutzer fehlten. Das hat sich nun mit Entwicklungen wie Programmatic Advertising und Real Time Bidding geändert. Basierend auf der Auswertung des anonymisierten Nutzungsverhaltens von Online-Nutzern können mittlerweile in Echtzeit Gebote auf einzelne Werbeeinblendungen abgegeben werden, ohne dass die beworbene Person dem Werbenden bekannt ist. Die Auswertung des Nutzungsverhaltens einzelner anonymer eindeutiger Schlüsselwerte wie Cookies im Browser des Endnutzers oder Advertiser-IDs in Apps auf Smartphones erlaubt die bestmögliche Platzierung von Werbung für den jeweiligen Nutzer im entsprechenden Nutzungskontext.

Die anonymisierte, datenschutzkonforme Auswertung von ca. 100 Milliarden Werbeeinblendungen am Tag führen spezialisierte Anbieter für werbende Unternehmen durch. Mit maschineller Mustererkennung lässt sich der zugrundeliegende Algorithmus für die Abgabe von Geboten fortwährend verbessern. Das Echtzeit-Bietsytem lernt selbständig. So wird sichergestellt, dass zum Beispiel die Werbung für ein neues Automodell auch einen Nutzer erreicht, der vermutlich gerade Interesse an einem Fahrzeugkauf hat. Die Vorteile sowohl für den Verbraucher, als auch für den Werbenden liegen auf der Hand. Der Verbraucher wird von unsinniger, breit gestreuter Werbung ohne Bezug zu seinen Präferenzen verschont, während der Werbende die richtigen Nutzer zur richtigen Zeit anspricht.

Werbetreibende sind bereit, für zielgerichtete Werbung an die passende Zielgruppe höhere Preise zu zahlen. So entsteht eine bessere Grundlage für den nachhaltigen Betrieb von Online-Services ohne Zusatzkosten für den Verbraucher.

## 4.12 Industrie 4.0

Big-Data-Technologien leisten einen unverzichtbaren Beitrag für Industrie 4.0.<sup>66</sup> Sie beeinflussen insbesondere die vier Bereiche Smart Factory, Smart Operations, Smart Products und Smart Service.

---

66 Vgl. (Roland Berger/BDI, 2015) sowie (McKinsey, 2015). Die McKinsey-Studie identifiziert fünf Handlungsfelder, auf denen Industrieunternehmen nun aktiv werden sollten: Handlungsfeld 1 lautet: »Daten besser nutzen«: Unternehmen sollten die komplette Wertschöpfungskette und den gesamten Lebenszyklus eines Produkts digital abbilden. Bisher wird nur rund ein Prozent der in der Produktion anfallenden Daten genutzt – ein enormes Potenzial liegt damit brach.

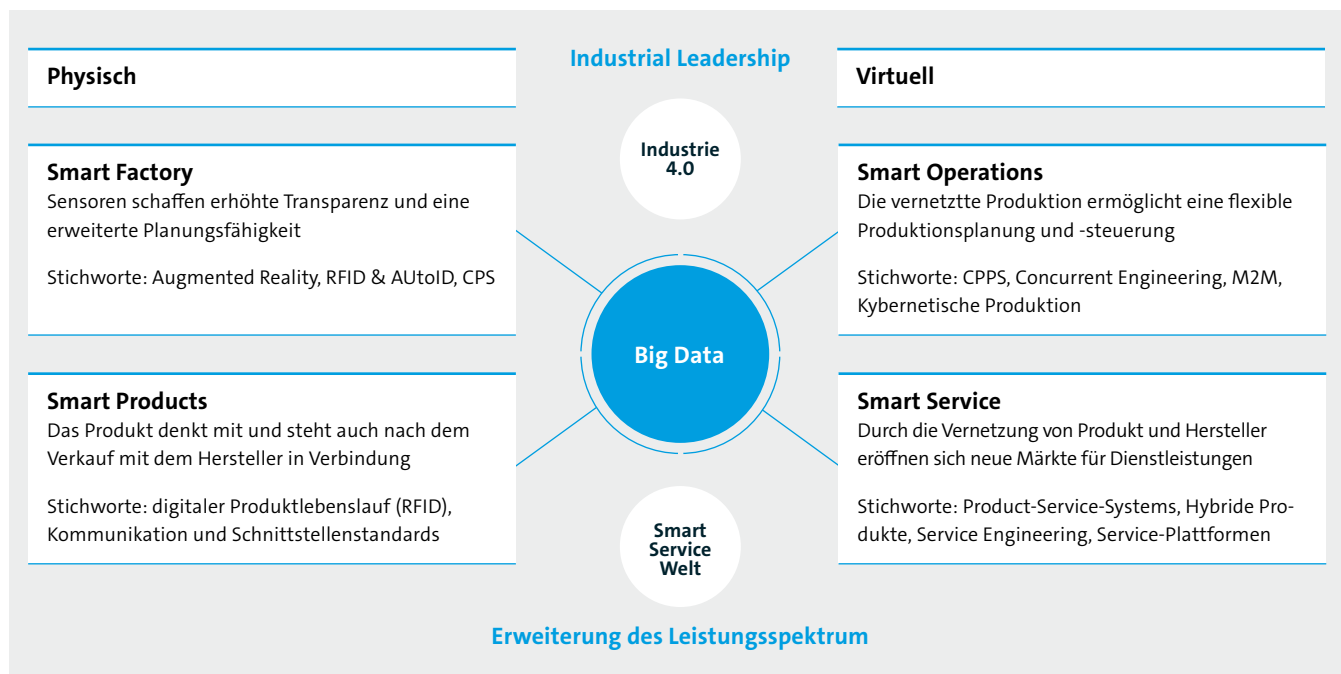


Abbildung 1: Big Data als Motor für Industrie 4.0<sup>67</sup>

Big Data gilt als eine der Triebkräfte für die vierte industrielle Revolution (Industrie 4.0). Die Möglichkeit, große und inhomogene Datenmengen zu erfassen und praktisch in Echtzeit zu verarbeiten, lässt die Digitalisierung der Industrie in absehbarer Zukunft zur Realität werden.

Der Einfluss von Big Data wird in vier Bereichen<sup>68</sup> evident (vgl. Abbildung 1). Zum einen ermöglicht es flexible und intelligente Fabriken (Smart Factory, Smart Operations), zum anderen äußert sich der Einfluss in den damit verbundenen qualitativ neuen Produkten (Smart Products) und Dienstleistungen (Smart Service) (vgl. Tabelle 2).

**Tabelle 2: Einflussbereiche von Big Data**

Einflussbereich	Erläuterung
<b>Smart Factory</b>	Die Ausstattung der Fabriken mit einer Vielzahl von Sensoren steigert zunächst die akquirierte Datenmenge und bei richtiger Verarbeitung auch die Transparenz. Ein Ziel dabei ist, die Realität der Produktion möglichst mit einem scharfen digitalen Schatten abzubilden, um eine möglichst genaue Produktionsplanung und folglich auch eine hohe Liefertermintreue zu gewährleisten. Mit einer Planungssicherheit im Rücken lassen sich Kunden dazu einladen, die Fertigung des gekauften Produktes (z. B. Auto, Motorrad, ...) unmittelbar mitzuerleben.

67 Quelle: FIR an der RWTH Aachen.

68 Die Deutsche Akademie der Technikwissenschaften treibt die vier Bereiche mit ihren umfassenden Initiativen »Industrie 4.0« und »Smart Service Welt« voran.

Einflussbereich	Erläuterung
<b>Smart Operations</b>	Eine Ausstattung der Produktion mit Sensortechnik eröffnet zudem neue Arten, die Produktion virtuell zu vernetzen. Dadurch lässt sich die Flexibilität steigern. Und weitere Einflussfaktoren, wie beispielsweise Energieverbrauch (und Anbindung an das Smart Grid), können berücksichtigt werden. Mit der Zustandsüberwachung von Maschinen wird frühzeitig vor bevorstehenden Ausfällen gewarnt. Bei unvorhergesehenen Ausfällen kann das vernetzte Produktionssystem schnell und flexibel mit Auftragsverschiebung reagieren, ohne dabei die Einflussfaktoren der Zielvorgabe außer Acht zu lassen.
<b>Smart Products</b>	Neben der intelligenten Fabrik wird auch das Produkt mit Intelligenz ausgestattet – beispielsweise über die Anbindung an eine Datenplattform. Das Verlassen des Firmengeländes unterbricht nicht zwangsweise den Kontakt zum Hersteller. Vielmehr erhält der Hersteller weitere wertvolle Daten und Informationen über den Einsatz. Mithilfe von intelligenten Auswertungen können Produkte besser auf die Kundenwünsche zugeschnitten werden. Und häufig auftretende Fehler helfen der Entwicklungsabteilung bei der Produktoptimierung.
<b>Smart Service</b>	Die neue Intelligenz sowie die Verfügbarkeit von Daten über die Produkte und ihre Nutzer lassen ergänzende Dienstleistungen oder sogar komplett neue Geschäftsmodelle entstehen <sup>69</sup> . Denkbar wäre beispielsweise, dass Kunden teure Maschinen nicht mehr selbst erwerben, sondern ihre Funktionalität »as a service« in Anspruch nehmen. Der Hersteller würde dann die Maschine zur Verfügung stellen, sich um ihre Funktionsfähigkeit kümmern und einen vertraglich fixierten Betrag pro produzierter Einheit erhalten. Eine Auswertung der gewonnenen Daten kann zudem zu weiteren Optimierungen der Maschine sowie der anknüpfenden Services führen.

Insgesamt wird deutlich, dass Big-Data-Technologien einen unverzichtbaren Beitrag für Industrie 4.0 leisten.

## 4.13 Telekommunikation

Mit Big-Data-Technologien erhöhen die Telekommunikationsunternehmen die Leistungsfähigkeit ihrer Infrastruktur. Sie können außerdem anonyme Bewegungsmuster ihrer Kunden erzeugen, die bei der Verkehrsplanung und Steuerung auf allen Ebenen sowie bei der Bewältigung zahlreicher weiterer, gesellschaftlich relevanter Problemstellungen hilfreich sind.

Der Mehrwert von Big-Data-Analysen für die Telekommunikation ergibt sich im Wesentlichen aus zwei Wertschöpfungsmodellen:

- Das erste Modell verbessert den Geschäftsbetrieb von Telekommunikationsunternehmen. So lässt sich z. B. die Leistungsfähigkeit der Telekommunikationsinfrastruktur unter Verwendung von Big Data signifikant steigern.
- Beim zweiten Modell geht es um die Optimierung von Wertschöpfungsprozessen bei Geschäftskunden und um die Steigerung gesellschaftlich wünschenswerter Effekte.

<sup>69</sup> Vgl. (Bitkom, 2015).

Dynamische Analytics-Lösungen nutzen die Fähigkeit von Telekommunikationsunternehmens, Bewegungsmuster zu erstellen, die auf Lokationsdaten der Teilnehmer sowie auf CRM-Daten basieren und mittels statistischer Methoden Bewegungsströme darstellen können. Verkehrsbetreiber<sup>70</sup> erhalten so wichtige Informationen für die Bedarfsplanung und Steuerung.

## 4.14 Verbesserung des Datenschutzes in Rechenzentren

Big-Data-Ansätze sichern die Überwachung der Einhaltung von Datenschutzrichtlinien in Rechenzentren.

Was passiert mit Daten, wenn ein Kunde sie einem Unternehmen übermittelt?

Die übermittelten Kundendaten werden für zukünftige Verarbeitungsschritte erst einmal gespeichert. Die Speicherung erfolgt meist zentral in einem Rechenzentrum – zusammen mit den Daten anderer Kunden, um diese effizient über unterschiedlichste Abfragemöglichkeiten in Zusammenhang zu bringen.

Innerhalb eines Rechenzentrums sind viele Mitarbeiter für den ordnungsgemäßen Betrieb zuständig. Um sicherzustellen, dass sie nur die ihnen übertragenen Tätigkeiten mit den übermittelten Daten durchführen, muss jede mögliche Interaktion überwacht werden. Die Interaktionen variieren zwischen genau definierten bis hin zu völlig freien Zugriffsmöglichkeiten. Je weniger definiert ein Zugriff stattfindet, desto weniger strukturiert werden die Überwachungsdaten abgelegt. Dies kann sich von Text- über Audio- wie auch Videodaten erstrecken.

Das soll anhand realistischer Zahlen eines IT-Providers für Finanzdienstleister verdeutlicht werden: Das Unternehmen hat 3000 Mitarbeiter, ebenso viele Arbeitsplätze, 150 unterschiedliche Tools (wie z. B. für Datenbanken, Auswertungen, Betriebssysteme, etc.), 1000 Server und viele der Mitarbeiter haben einen VPN-Zugang, der einen 7x24h-Zugriff ermöglicht. Jeder Mitarbeiter hat ein bestimmtes Benutzerprofil, über das seine Zugriffsmöglichkeiten so beschränkt sind, dass er seine alltägliche Arbeit durchführen kann. Alle Interaktionen aller Mitarbeiter mit allen oben genannten Elementen werden geloggt – dies beinhaltet auch die daraus resultierenden Ergebnisse. Diese Daten werden historisiert, um bei einer Meldung bzw. automatisierten Erkennung eines Vorfalles alle Beweisdaten zu haben.

Die Herausforderung besteht nun darin, aus dieser großen, unstrukturierten Datenmenge zu erkennen, welche Operationen die Benutzer mit den Daten vorgenommen haben und ob sie dazu berechtigt waren.

Berechtigt ist nur, wer einen Auftrag hat. Ein Auftrag kann ein Projekt sein oder aber auch eine Fehler- bzw. Wartungssituation. Der Zusammenhang zwischen einer durchgeführten Tätigkeit

---

<sup>70</sup> Vgl. Abschnitt 4.2 und Abschnitt 4.10.

mit Datenzugriff und einem Auftrag muss gegeben sein. Ansonsten hat der entsprechende Mitarbeiter mit arbeitsrechtlichen Konsequenzen und im schlimmsten Fall mit straf- oder zivilrechtlichen Verfahren zu rechnen.

Auch mit der heutigen Technologie ist es nicht möglich, eine 100-prozentige Vollständigkeit der Überwachung zu gewährleisten. Deshalb werden manuelle Prüfungen durchgeführt, die über einen Zufallsgenerator ausgewählt werden.

Bei der Überwachung der Einhaltung von Datenschutzrichtlinien wird Big Data im Interesse der Kunden eingesetzt.







# 5 Bisher ungenutzte Chancen beim Einsatz von Big Data

## 5 Bisher ungenutzte Chancen beim Einsatz von Big Data

In anderen Ländern werden Big-Data-Anwendungen für Aufgaben eingesetzt, bei denen in Deutschland besondere Rahmenbedingungen gelten, die bisher dem Einsatz entgegenstanden. Es steht die Frage im Raum, ob diese Rahmenbedingungen für alle Ewigkeit unverrückbar bleiben müssen, wenn sie sich als Hemmschuh im globalen Wettbewerb der Wirtschaftsstandorte erweisen.

### Mehr Transparenz – mehr Akzeptanz

Jeder neuen Technologie schlägt Skepsis entgegen, schließlich ging es ja vorher auch ohne. Zudem stellt sich leicht ein Unbehagen gegenüber dem Unbekannten ein. Im Bereich Big Data gilt das in besonderem Maße. Alle Bedenken sind ernst zu nehmen und können in der Diskussion letztlich nur förderlich sein. Mehr Transparenz<sup>71</sup> kann zu mehr Akzeptanz führen und schließlich dabei helfen, die positiven Potenziale von Big Data zu heben – und damit sind nicht nur wirtschaftliche Vorteile einzelner Unternehmen gemeint, sondern auch allgemeine Produktivitätssteigerungen und gesamtgesellschaftliche Fortschritte.

### Öffentliche Verwaltung

Im Bereich der Steuereinnahmen herrscht in Deutschland eine föderale Zuständigkeit, die zu unterschiedlichen Auslegungsvorschriften je nach Bundesland führen. Wichtiger noch als die subjektiv empfundenen Ungleichbehandlungen, etwa in der Verfolgung von einzelnen Steuersäumigen, sind jedoch kriminelle Ansätze, die auf die fehlende Gesamtsicht setzen. Sogenannte Umsatzsteuerkarusselle lassen sich durch eine geschickte Konstruktion so leichter als in anderen europäischen Ländern aufbauen. Durch ein Netz von verteilten Firmen werden Vorsteuer und Steuererstattungen so verschachtelt und verzahnt, dass am Ende dem Staat nicht nur Steuern entgehen, sondern womöglich sogar mehr Steuern erstattet werden, als in einem entsprechenden Geflecht überhaupt gezahlt wurden. Eine komplette Datenerhebung aller steuerlichen Sachverhalte ließe sich sehr viel einfacher mittels analytischer Verfahren auf verdächtige Muster und Zusammenhänge hin untersuchen. In Deutschland ist der Aufwand ungleich höher – die Schlupflöcher mithin größer.

### Soziales

Eine der größten Errungenschaften des modernen Staats sind seine Sozialleistungen. Mit Hilfe analytischer Methoden können der Missbrauch von Sozialleistungen verringert und Ausgaben vermieden werden. So könnte es gelingen, durch ein aktives und fokussiertes Eingreifen, zum Beispiel durch gezielte Fördermaßnahmen oder spezifische Bildungsangebote bereits im Vorfeld einschlägige »Karrieren« zu verhindern. Jeder erfahrene Sozialarbeiter kennt die maßgeblichen

---

71 Vgl. dazu Abschnitt 7.3 und 7.4.

auslösenden Faktoren – mit Big Data Analytics ließen sich bereits deren erste Anzeichen erkennen. Was in Frankreich und anderswo möglich ist, scheitert in Deutschland am Datenschutz (vgl. Leitlinie 12). Übrigens gilt das genauso wie der Zeitverzug beim Bekanntwerden von Missbrauchsfällen von Leistungsbezug. Wären hier mehr Systeme integriert und mit automatischen Alertfunktionen ausgestattet, ließe sich so manche Auszahlung vermeiden.

## Gesundheit

Der Bereich Gesundheit ist höchst sensibel und verdient besonderen Schutz. Umgekehrt erwartet jeder Patient die bestmögliche Behandlung in seinem konkreten Fall. Nun gibt es gerade in diesem Bereiche eine Unmenge von Daten: Verschreibungsinformationen, Diagnosen, Verlaufsprotokolle, Medikamentenwirksamkeitsstudien und so weiter. Diese Daten »gehören« jedoch unterschiedlichen Institutionen und können deshalb auch nicht so einfach gemeinsam ausgewertet werden. Wäre dies möglich, könnten anhand der Millionen Datenpunkte, die etwa im Bereich des Morbi RSA<sup>72</sup> anfallen, funktionierende von nicht funktionierenden Behandlungsmustern getrennt werden (vgl. Leitlinie 7 und Leitlinie 10). Zumindest ließe sich mit einer signifikanten Wahrscheinlichkeit sagen, welche Behandlung »normalerweise« Erfolg verspricht und welche »fast nie« wirkt. Welche Entscheidung der Arzt in seinem konkreten Fall trifft, dürfte selbstverständlich nicht determiniert werden. Aber wären nicht alle Patienten dankbar, von den Erfahrungen »aller« anderen vergleichbaren Fälle zu profitieren (vgl. Leitlinie 12)?

## Hinderungsgründe

Fasst man diese und weitere Beispiele zusammen, ergeben sich zwei wesentliche Hinderungsgründe:

- Zweckbindung<sup>73</sup> im Bundesdatenschutzgesetz<sup>74</sup>:  
Der Reiz und der Mehrwert vieler Big-Data-Projekte entsteht im explorativen Erforschen und Kombinieren bisher getrennt erfasster Einzeldatenbestände. Damit lassen sich Fragestellungen nach dem Zusammenhang von Wetter (Temperaturdaten) und Abverkauf von Eiscreme beantworten. Bei personenbezogenen Daten fehlt (bisher) oft ein gesetzlicher Erlaubnistatbestand, der eine nachträgliche Verwendung für andere Zwecke, als ausdrücklich bei der Erfassung abgefragt, ohne die ausdrückliche Einwilligung legitimiert (vgl. Leitlinie 4). Das führt dazu, dass bestehende Datenbestände in vielen Fällen nur anonymisiert oder mit einer neuen Einwilligung ausgewertet werden können. Beides ist in der Praxis oft nur schwer zu leisten.
- Starke Anforderungen an Anonymisierung<sup>75</sup>:  
Der übliche und häufig verfolgte Ausweg bei nicht ausreichenden Einwilligungen ist die

72 Risikostrukturausgleich der gesetzlichen Krankenkassen.

73 Vgl. dazu Abschnitt 9.

74 Vgl. [http://www.gesetze-im-internet.de/bdsg\\_1990](http://www.gesetze-im-internet.de/bdsg_1990) (abgerufen am 31. August 2015).

75 Vgl. dazu Abschnitt 9.

Anonymisierung von personenbezogenen Daten. Dann sind zwar im Anschluss auch keine personenbezogenen Handlungen im engeren Sinne möglich, wohl aber lassen sich Verbraucheransprachen zusammenfassen und optimieren. Rechtlich ist allerdings in Deutschland schon nicht zweifelsfrei geklärt, ob man vorhandene Daten ohne weiteres für eine weitere anonyme Nutzung anonymisieren darf oder ob für den Vorgang der Anonymisierung eine eigene Rechtsgrundlage benötigt wird. Technisch lassen sich Anonymisierungen auf verschiedene Arten realisieren. Wenn die Fallzahl von nachträglich gebildeten Gruppen zu klein werden, lassen sich unter Umständen wieder reale Personen zuordnen, so dass die Daten nach den geltenden rechtlichen Anforderungen allenfalls als pseudonymisiert, aber nicht als anonymisiert gelten. Die Verarbeitung von pseudonymen Daten wird nach geltendem Datenschutzrecht derzeit kaum privilegiert, obwohl die Pseudonymisierung in Kombination mit technischen und organisatorischen Maßnahmen (zum Schutz vor De-Identifizierung) resultierende Risiken für betroffene Personen maßgeblich reduziert.<sup>76</sup>

## Lösungsansätze

- Woran es derzeit in Deutschland mangelt, sind vertrauenswürdige Organisationen (Daten-Treuhänder), die Datenbestände von zwei Parteien aufnehmen und jeder Partei das »Gold« aus den Daten (anonymisiert) zurückgibt, um daraus neue digitale Geschäftsmodelle für beide Seiten zu erstellen<sup>77</sup> (vgl. Leitlinie 12).
- Auch fehlt es an erprobten Technologien, die diese Anonymisierung und Pseudonymisierung so einfach machen, dass es für alle Beteiligten transparent und handhabbar bleibt. Erfolversprechende Ansätze beziehen auch Hardwarekomponenten mit ein oder arbeiten mit ausgefeilten Zertifikathierarchien.
- Technische Leitlinien und anerkannte Standards zur Anonymisierung und Pseudonymisierung sind (weiter) zu entwickeln, um einen Maßstab für alle Beteiligten zu schaffen, an der die Qualität der Anonymisierung und Pseudonymisierung rechtssicher gemessen werden kann.
- Die wichtigste Errungenschaft allerdings wäre ein modernes Datenschutzrecht, das den tatsächlichen Gepflogenheiten der Gegenwart Rechnung trägt und einen möglichst global einheitlichen Raum definiert (vgl. Leitlinie 12 und Kapitel 9).

<sup>76</sup> Ein Beispiel soll das verdeutlichen: In der Erforschung von bisher nicht entdeckten Nebenwirkungen von Arzneimitteln lassen sich anonymisierte Daten verwenden. Auf diese Dienstleistung spezialisierte Anbieter aggregieren solche Daten und durchforsten sie nach Zusammenhängen. Hier kommt es immer wieder zu Rechtsstreitigkeiten bei der Frage der Qualität der Anonymisierung. Oft ist die Möglichkeit der De-Identifizierung bei derartigen Projekten im Interesse des Patienten auch gewünscht, da Erkenntnisse der Forschung u. a. für die Therapie des Patienten verwendet werden können. Auch hier kann die erleichterte Nutzung von pseudonymen Daten Lösungsmöglichkeiten eröffnen.

<sup>77</sup> Diesen Service bietet als einer der wenigen Anbieter die Seculytics GmbH. Der Lösungsansatz besteht im Privacy-Preserving Data Mining mit strikt getrennten Datentöpfen.







# 6 Bedeutung der Daten- wirtschaft in der zukünftigen Wirtschaft

# 6 Bedeutung der Datenwirtschaft in der zukünftigen Wirtschaft

Die Datenwirtschaft wird die Geschäftsmodelle vieler Branchen unserer Industrie- und Dienstleistungslandschaft umkrempeln. Zahlreiche Beispiele aus Unternehmen zeigen die Entstehung bzw. Modifizierung von Geschäftsmodellen durch Big Data auf. Big Data ist unternehmerische Realität und eröffnet neue Möglichkeiten für innovative Unternehmen und Startups. Der Wert eines Produktes aus Verbrauchersicht wird immer stärker durch intelligente Software-Funktionalität und kontextsensitive Bereitstellung von Daten geprägt.

## Technologie- und Strategietrends

»Big Data – Die Daten werden zum Produkt« – dieser Merksatz beschreibt in Kurzform die Auswirkungen des Phänomens der Datenwirtschaft auf die Wertschöpfungs- und Geschäftsmodelle in der Industrie- und Dienstleistungslandschaft. Was sich in den letzten Jahren als strategischer Trend abzeichnete, beginnt sich heute als unternehmerische Realität zu manifestieren und verändert nachhaltig die Wettbewerbslogik in vielen Märkten.

Es lassen sich derzeit drei grundlegende Trends ausmachen, deren Konvergenz in den kommenden drei bis fünf Jahren für ein extrem beschleunigtes Wachstum der »Data Economy« sorgen werden:

- Big Data und neue Analytics-Verfahren,
- IoT und Software-defined Products,
- Software-defined Business.

In dem genannten Zeitraum wird sich auch der Abstand zwischen den klassisch organisierten und den datengetriebenen Unternehmen weiter vergrößern. Viele etablierte Spieler werden vom Markt verschwinden, da ihre Marken in der digitalen Welt nicht mehr sichtbar und die Geschäftsmodelle ihrer Produkte nicht mehr kompatibel mit den digitalen Kommunikations- und Kommerzialisierungsströmen sind. So sind aktuell die Banken und Versicherungen in Alarmbereitschaft, da sich sogenannte FinTech-Start-ups mit ihren anwenderfreundlichen und mobilen Lösungen in die Finanzströme einklinken und den etablierten Finanzkonzernen die jungen und zahlungsfreudigen Verbraucher abjagen.

## Big Data und neue Analytics-Verfahren

Ohne die neue Generation an Datenbanktechnologien, Analytics-Verfahren und hochskalierenden IT-Infrastrukturen á la Cloud Computing ist eine Datenwirtschaft nicht denkbar. Die in den letzten Jahren als Big Data umschriebenen neuen Technologien und Analysewerkzeuge stellen die Grundlage der neuen datengetriebenen Geschäftsmodelle und Prozesse dar. Auch hier konvergieren derzeit verschiedene Technologien und schaffen vollständig neue Möglichkeiten für innovative Unternehmen und Start-ups. So lassen sich heute Machine-Learning-Algorithmen auf Cloud-Plattformen »as a Service« anwenden und stundengenau abrechnen. Auch Anwen-

dungen im Kontext Cognitive Computing machen derzeit große Fortschritte, da die Prognosegüte dank steigender Datenvolumina und unbegrenzter Rechenleistung in der Cloud immer weiter zunimmt.

## IoT und Software-defined Products

Mit dem Internet der Dinge beginnt eine engmaschige Verzahnung von analoger und digitaler Welt. Angetrieben durch kostengünstige Sensorik werden immer mehr Maschinen und Anlagen vernetzt. So werden sowohl Industrie- als auch Dienstleistungsprozesse granular analysierbar, optimierbar und sogar »programmierbar«. Gleiches gilt zukünftig auch im Kontext von Produktentwicklung und Produktdesign von Haushaltsgeräten, Konsumelektronik, Kleidung oder auch Autos. Der Wert eines Produktes aus Verbrauchersicht wird immer stärker durch intelligente Software-Funktionalität und kontext-sensitive Bereitstellung von Daten geprägt. Das gilt nicht nur für Premium-Automobile, wie den Tesla Model S, sondern auch für Skibrillen mit eingebautem Heads-up-Display, welches GPS-Daten und WLAN-Services integriert, sowie für Bohrmaschinen und Rasenmäher. Die Einzigartigkeit des Produkts ist durch Software definiert und die Bereitstellung und Verarbeitung von produkt- und kontextrelevanten Daten spielt für den Anwender eine immer zentralere Rolle. Sie bestimmt somit maßgeblich den Differenzierungsfaktor eines Produktes sowie die Zahlungsbereitschaft der Verbraucher.

## Software-defined Business

Auf dieser Basis entstehen softwaregetriebene Geschäftsmodelle, bei denen die physische Beschaffenheit (Material, Design, Ergonomie etc.) der Produkte in den Hintergrund tritt und dem digitalen Funktions- und Kommunikationsspektrum des Produktes Platz macht. Diese Transformation findet sukzessiv statt und entkoppelt in nahezu allen Branchen die physischen von den digitalen Wertschöpfungsketten. So lagern schon heute moderne Konzerne einen Großteil ihrer Prozesse an Partner aus und »managen« lediglich die Produktentwicklung und die Verbraucherbeziehung selbst. Erfolgreiche Beispiele sind Apple oder Nike. In der Welt des Software-defined Business wird vor allem das Management der digitalen Verbraucherbeziehung (»Digital Customer Experience«) zum Erfolgsfaktor. Wer die Schnittstelle zum Verbraucher beherrscht und dessen Daten, Gewohnheiten und Präferenzen in Vertrieb und Support zu handhaben weiß, zählt zu den Gewinnern. Vorgemacht haben dies Unternehmen wie Netflix oder auch Zalando, die sich erfolgreich zwischen Produzent und Verbraucher positioniert haben und mittels intelligentem Verbraucherdatenmanagement und Analytics die Zahlungsbereitschaft voll ausschöpfen und mit sehr genau designten Angeboten punkten.

## Data Economy – Erfolgsfaktoren und Einstiegsoptionen

Diese drei beschriebenen Technologie- und Strategietrends verstärken sich wechselseitig. Auch erfassen sie nahezu alle Branchen, vom Handel, über die Energieerzeuger bis hin zur verarbeitenden Industrie, wenn auch mit unterschiedlicher Geschwindigkeit. Doch wo liegen die Erfolgsfaktoren und Einstiegsoptionen in die schöne, neue »Data Economy«? Welche Fähigkeiten brauchen Unternehmen, um in der Datenwirtschaft zu den Marktführern zu zählen?

Es zählen insbesondere vier Bereiche:

- Inhaberschaft/Besitz und Verwertungsrechte an Datenbeständen,
- Fähigkeiten zur Analyse und Aufarbeitung von Daten,
- Exzellenz im Bereich der Gestaltung digitaler Produkte und einer »Digital Customer Experience«,
- Vertrauen.

Inhaberschaft/Besitz und Verwertungsrechte an Datenbeständen bilden für Unternehmen eine gute Ausgangsbasis für den Start in die Datenwirtschaft. Doch muss im Unternehmen eine entsprechende Awareness und Datenkultur vorherrschen, so dass der potenzielle Wert der Datenbestände erkannt und die Kommerzialisierungsmöglichkeiten richtig eingeschätzt werden können.

Ohne interne Fähigkeiten zur Analyse und Aufarbeitung der Daten (Analytics Capabilities) und eine Mannschaft aus Data Scientists liegen die Daten in vielen Unternehmen meist brach. Allerdings lassen sich diese Fähigkeiten zumindest teilweise von Beratern und Dienstleistern einkaufen.

Exzellenz im Bereich der Gestaltung digitaler Produkte und einer »Digital Customer Experience« lässt sich dagegen nur schwer extern zukaufen. Hier liegt einer der zentralen Wettbewerbsfaktoren digitaler Champions bzw. Leader der Data Economy.

Einen Vorteil, den viele etablierte Unternehmen mit an den Tisch bringen, ist das Vertrauen einer breiten Kundenbasis. Als »Trust Provider« können etablierte Brands auch in der Datenwirtschaft reüssieren, indem sie für innovative Start-ups und Geschäftsmodelle die Basis für den Markteintritt liefern oder als Verbraucherschnittstelle agieren.

Im Jahr 2016 werden weltweit schon über 140 Milliarden Euro Umsatz mit datengetriebenen Produkten und Geschäftsmodellen generiert.<sup>78</sup> Die Tendenz ist stark steigend. Die Datenwirtschaft ist also schon Realität und wird in den kommenden drei bis fünf Jahren auch eine neue Generation an Managern an die Spitze großer Konzerne befördern.

---

78 Vgl. (Crisp Research, 2014).





# 7 Fragestellungen für Big-Data-Leitlinien

# 7 Fragestellungen für Big-Data-Leitlinien

Der Einsatz von Big-Data-Technologien ist Gegenstand eines intensiven öffentlichen Diskurses. Im Kapitel 7 werden in zehn Abschnitten ausgewählte Fragen aufgegriffen, die aus Sicht des Bitkom besonders wichtig sind.

- **Rechtssicherheit auf europäischer Ebene:** Für Big Data in der internationalen Dimension sind einheitliche, international geltende Regeln erforderlich. Diese müssen für die an der Datenverarbeitung Beteiligten wie auch für die Betroffenen ein hohes Maß an Rechtssicherheit, Transparenz und Vertrauen schaffen. Dies soll mit der in Diskussion befindlichen EU-Datenschutz-Grundverordnung erreicht werden (vgl. Abschnitt 7.1). Die EU-Datenschutz-Grundverordnung bleibt derzeit aber noch hinter den Möglichkeiten und Erfordernissen der Digitalisierung zurück. Um für Industrie 4.0, Big Data, Cognitive Computing und andere Innovationen günstige Rahmenbedingungen zu schaffen, sollte ein politischer Prozess zur Weiterentwicklung und praktischen Umsetzung der Verordnung aufgesetzt werden (vgl. Kapitel 9).
- **Verbraucherakzeptanz von Big Data:** Damit die Verbraucher Big-Data-Anwendungen akzeptieren, müssen ihre Vorbehalte und Ängste verstanden und ernst genommen werden. Verbraucher benötigen mehr Unterstützung, um ein Grundverständnis für die eingesetzten Technologien (vgl. Abschnitt 7.2) und die Arbeit von Data Scientists (vgl. Abschnitt 7.3) zu entwickeln. Die Akzeptanz von Big-Data-Technologien in der breiten Öffentlichkeit wird deutlich zunehmen, wenn sich die Technologieanbieter zu einer ethisch fundierten Herangehensweise bekennen (vgl. Abschnitt 7.5).
- **Transparenz von Big-Data-Anwendungen:** Transparenz bei Big-Data-Anwendungen bedeutet, dass die Verarbeitung der Daten mit zumutbarem Aufwand von Betroffenen nachvollzogen, überprüft und auch bewertet werden kann. So verstanden, ist Transparenz ein wesentlicher Schlüssel für die Akzeptanz von Big-Data-Lösungen in der öffentlichen Meinungsbildung (vgl. Abschnitt 7.3).
- **Individualisierte Versicherungstarife und Solidarprinzip:** Individualisierte Versicherungstarife müssen nicht im Widerspruch zum Solidarprinzip stehen. Sie eröffnen die Möglichkeit, ein gesellschaftlich akzeptiertes Verhalten zu fördern, ohne die Risiken völlig auf das Individuum abzuwälzen. Big-Data-Analysen können auch einen digitalisierten Schutz ermöglichen, indem auf ihrer Grundlage Empfehlungen generiert werden. Verbraucher sollten zwischen der Ablehnung einer individuellen Vermessung ihrer Gewohnheiten bis hin zur Akzeptanz und freiwilligen Datenlieferung wählen können (vgl. Abschnitt 7.6).
- **Bedeutung automatisierter Datensammlung:** Die automatisierte Sammlung von Daten bietet sowohl für den Verbraucher als auch für Unternehmen einen erheblichen Mehrwert. Zu gewährleisten ist dabei aber, dass die rechtlichen Vorgaben beachtet werden und der Verbraucher



den Überblick darüber behält, welche Daten seine Geräte übermitteln, um ihm selbstbestimmtes Handeln zu ermöglichen (vgl. Abschnitt 7.7).

- **Recht auf Vergessenwerden:** Ein »Recht auf Vergessenwerden« einklagen zu können, geht an der Realität vorbei. Vielmehr geht es um selbstbestimmtes Löschen und darum, überhaupt zu erfahren, wo welche personenbezogenen Daten liegen (vgl. Abschnitt 7.8).
- **Grenzen für Prognosen:** Prognosen auf der Basis von Big-Data-Einsatz können an Grenzen geraten: Sie dürfen nicht die Chancengleichheit von Menschen beschneiden (vgl. Abschnitt 7.9).
- **Analyse Sozialer Netzwerke:** Unternehmen sollten bei der Analyse der Kommunikation in Sozialen Netzwerken umsichtig agieren und die Interessen der Betroffenen angemessen berücksichtigen (vgl. Abschnitt 7.10).

## 7.1 Big Data im internationalen Wettbewerb

Für Big Data in der internationalen Dimension sind einheitliche, international geltende Regeln erforderlich. Diese müssen für die an der Datenverarbeitung Beteiligten wie auch für die Betroffenen ein hohes Maß an Rechtssicherheit, Transparenz und Vertrauen schaffen. Dies soll mit der EU-Datenschutz-Grundverordnung erreicht werden.

### Noch erhebliches Maß an Rechtsunsicherheit

Die internationale Dimension von Big Data ist immer dann relevant, wenn Daten aus einem oder mehreren Ländern in einem oder mehreren anderen Ländern verarbeitet werden.<sup>79</sup> Aus Sicht des Datenschutzes ist in diesen Fällen von besonderem Interesse, welche Regeln zur Anwendung kommen. In der Europäischen Union ist das auf der Richtlinie 95/46 EG basierende Datenschutzrecht weitgehend harmonisiert. Darauf aufbauend sind die Anforderungen an die Datenverarbeitung innerhalb der Europäischen Union an die Verarbeitung im Inland angeglichen. Ebenfalls in der Richtlinie geregelt sind die Fälle, in denen Drittländer außerhalb der Europäischen Union an der Datenverarbeitung beteiligt sind. Dies legt zunächst einen klaren Regelungsrahmen nahe, der für alle Beteiligten die benötigte Rechtssicherheit in der Datenverarbeitung bietet. Allerdings bestehen aufgrund der Umsetzung der Richtlinie 95/46 EG in den Mitgliedsstaaten graduelle Unterschiede im Datenschutzrecht, mit der Folge von Auslegungs-

<sup>79</sup> Hofstetter weist darauf hin, dass verschiedene Sichten auf Big Data in den USA und im kontinentalen Europa die Unterschiede im Menschenbild und im Marktverständnis widerspiegeln (vgl. (Hofstetter, 2015, S. 36f.)). »Free trade als Letztbegründung für die Erfassung, die Speicherung und den Handel mit Big Data hebt die Rechte des Menschen auf eine selbstbestimmte Zukunft auf, indem sie ihn zum Gegenstand des Wirtschaftens verkürzt.« (Vgl. (Hofstetter, 2015, S. 37)) Und später: »Big Data spricht dem Menschen die Subjekteigenschaften ab und ordnet ihn der Technik unter – mit allen Konsequenzen für den Einzelnen und seine Rechte, für Wirtschaft, Politik und Gesellschaft. ... Das Subjekt und seine Freiheitsrechte erodieren, mit ihm die soziale Marktwirtschaft und, wie nebenbei, auch Ethik und Moral ökonomischen Handelns.«

streitigkeiten sowohl für den Datenverkehr innerhalb der EU als auch für den Datentransfer in Drittstaaten, also außerhalb der EU und des EWR. Statt der benötigten Rechtssicherheit besteht trotz der Harmonisierung durch die Richtlinie ein erhebliches Maß an Rechtsunsicherheit. Gerade vor dem Hintergrund einer zunehmenden globalisierten Datenverarbeitung sind aber möglichst einheitliche und verlässliche Regeln nicht nur wünschenswert, sondern auch erforderlich.

### Neue EU-Datenschutz-Grundverordnung soll Rechtsrahmen vereinheitlichen

Rechtssicherheit, Transparenz und Vertrauen sind maßgebliche Kriterien für eine Datenverarbeitung, nicht nur, aber insbesondere im Rahmen von Big Data. Auch wenn diese Kriterien gleichermaßen für nationale Big-Data-Anwendungen gelten, zeigt doch erst der internationale Kontext, in dem unterschiedliche Rechtsordnungen eine Rolle spielen, die Bedeutung dieser Kriterien. Nur wenn Rechtssicherheit durch ein möglichst klares und homogenes Recht besteht, können auch Transparenz und Vertrauen erzeugt werden.

In der Europäischen Union wird dieser einheitliche Rechtsraum in absehbarer Zukunft durch die Datenschutz-Grundverordnung hergestellt, die das auf der Richtlinie 95/46 basierende Datenschutzrecht ablösen wird. Wegen des Verordnungskarakters werden die Unterschiede reduziert, die sich aus der unterschiedlichen Umsetzung der Richtlinie 95/46 in nationales Recht ergeben haben und aus denen Rechtsunsicherheiten resultieren. Die Verordnung wirkt, im Gegensatz zu einer Richtlinie, unmittelbar für die jeweiligen Länder, d. h., es ist keine Umsetzung in das jeweilige nationale Recht erforderlich, so dass es keine Umsetzungsspielräume gibt. Die weiterhin mögliche unterschiedliche Handhabung von gleichen Sachverhalten durch die Aufsichtsbehörden muss durch Kohärenzverfahren zwischen den Aufsichtsbehörden beigelegt werden.

### Bedeutung der EU-Datenschutz-Grundverordnung für internationale Märkte

Aber auch für internationale, über das Territorium der EU hinausgehende Big-Data-Anwendungen trifft die Datenschutz-Grundverordnung mit dem in ihr verankerten Marktortprinzip eine wichtige Aussage. Zukünftig gilt europäisches Datenschutzrecht für alle Akteure auf dem europäischen Markt – unabhängig davon, ob Dienste oder Waren von innerhalb oder außerhalb der Europäischen Union angeboten werden. Damit besteht Klarheit und damit auch Rechtssicherheit über das anzuwendende Recht für alle an Big-Data-Datenverarbeitungsprozessen Beteiligten, egal, ob diese Daten innerhalb oder außerhalb der Europäischen Union verarbeiten.

Wie bereits im Abschnitt 5 angemerkt wurde, sollte ein modernes Datenschutzrecht im Big-Data-Kontext einen möglichst global einheitlichen Raum definieren. Mit der Datenschutz-Grundverordnung wird zwar keine globale Regelung geschaffen. Dennoch entsteht durch die klare Vorgabe auch für die Datenverarbeitung außerhalb der Europäischen Union ein beträchtlicher Rechtsraum mit gleichen Regeln. Um diesen Rechtsraum zu erhalten, muss sichergestellt werden, dass die durch die Datenschutz-Grundverordnung erreichte Rechtsvereinheitlichung und auch

das Datenschutzniveau in Bezug auf den europäischen Markt auch bei Sondervereinbarungen wie Safe Harbor oder Handelsabkommen beachtet und erhalten werden.

## Gleiche Wettbewerbsbedingungen für alle Marktteilnehmer

Mit einheitlichen Regeln für denselben Markt eröffnen sich zudem auch wirtschaftliche Perspektiven. Durch die Anwendung der europäischen Regeln auch auf Unternehmen außerhalb der Europäischen Union, die auf dem europäischen Markt agieren, wird ein Level Playing Field für alle Marktteilnehmer geschaffen, welches für fairen Wettbewerb sorgt. Gleichzeitig kann der Verbraucher sich leichter orientieren, wenn er weiß, dass alle Anbieter sich an die europäischen Regeln halten – egal ob sie ihre Leistungen aus Europa heraus erbringen oder aus anderen Staaten.<sup>80</sup>

Für Big Data in der internationalen Dimension sind einheitliche, international geltende Regeln erforderlich. Diese müssen für die an der Datenverarbeitung Beteiligten wie auch für die Betroffenen ein hohes Maß an Rechtssicherheit, Transparenz und Vertrauen schaffen. Dies sollte soweit wie möglich mit der Datenschutz-Grundverordnung erreicht werden.

## 7.2 Vertrauensentwicklung bei Verbrauchern – Best Practices

Damit die Verbraucher Big-Data-Anwendungen akzeptieren, müssen zwei Voraussetzungen erfüllt sein: Erstens gilt es, die Vorbehalte und Ängste der Verbraucher zu verstehen und ernst zu nehmen. Zweitens müssen den Verbrauchern die eingesetzten Technologien soweit nahegebracht werden, dass sie dafür ein Grundverständnis entwickeln können.

Big Data bietet für die Mehrheit der Unternehmen das Potenzial, Einnahmen, Marktanteile und Margen zu steigern. Begriffe wie Customer Journey oder optimales Marketinginvestment spiegeln dieses Potenzial wider. Grundlage hierfür ist die Nutzung großer Datenmengen, auch und insbesondere mit persönlichen Informationen. Diese Datenmengen werden mit dem Ziel analysiert, die Bedürfnisse aus den unterschiedlichsten Lebensbereichen der Verbraucher besser zu verstehen und mit spezifischen Angeboten darauf zu reagieren.<sup>81</sup>

Behörden und Unternehmen verfügen bereits über eine Vielzahl an Daten, deren Analysen einen Mehrwert für alle Beteiligten bedeuten könnten.

Neue Technologien wie Big Data führen oft zu einem Spannungsfeld zwischen dem technisch Möglichen und dem ethisch Vertretbaren. Dieses Spannungsfeld speist sich aus dem Antrieb und Wunsch des Menschen, Erleichterungen für seinen Alltag zu entwickeln (vgl. Kapitel 4), sowie aus der Skepsis gegenüber anstehenden gesellschaftlichen und technologischen Veränderungen.

<sup>80</sup> Vgl. (Kühling, 2015).

<sup>81</sup> Vgl. (Bitkom Research, 2015).

Der vom Fraunhofer-Institut für Sichere Informationstechnologie durchgeführte Bürgerdialog zeigte, dass die Akzeptanz für die Anwendung von Big Data dann am höchsten ist, wenn der Einsatz transparent dargestellt wurde und die Umsetzung von (Datenschutz-) Rechten erkennbar ist. Anwendungen aus den Bereichen Verbrechensbekämpfung, Wissenschaft und Medizin stießen im Bürgerdialog auf eine höhere Zustimmung als Anwendungen für Wirtschaft und Konsum.<sup>82</sup>

Kommerzielle Big-Data-Projekte werden in der Öffentlichkeit oft als intransparent wahrgenommen oder dargestellt.<sup>83</sup> Dieser Perzeption kann man mit vier Maßnahmen entgegenwirken:

- Verständnis für die Technologie fördern, z. B. durch die Erklärung von Anonymisierungsprozessen,
- Wahlmöglichkeit bei der Angabe von Daten anbieten, z. B. durch ein Verbraucherservice-Cockpit oder ein Kundenintegrations-Portal,
- Verständliche Informationen in den Datenschutzhinweise und -erklärungen bereitstellen,
- Risiken darstellen.

Die im Kapitel 8 formulierten Positionen und Leitlinien greifen diese Wege zu mehr Vertrauen bei den Verbrauchern in Big Data auf. Es gilt, die Verbraucher mit Nachhaltigkeit und »Wort halten« zu überzeugen, dass der kommerzielle Nutzen des Unternehmens keinen Nachteil für den Verbraucher bedeutet.

## 7.3 Transparenz bei Big-Data-Analysen

Transparenz bei Big-Data-Anwendungen bedeutet, dass die Verarbeitung der Daten mit zumutbarem Aufwand von Betroffenen nachvollzogen, überprüft und auch bewertet werden kann.

Transparenz bei Big-Data-Analysen und –Anwendungen (vgl. Leitlinie 2) bedeutet generell, dass die Verarbeitung der Daten mit zumutbarem Aufwand von Betroffenen nachvollzogen, überprüft und auch bewertet werden kann. Die Transparenz kann allen Beteiligten dienen, d. h. der verantwortlichen Stelle bei datenschutzrechtlichen Themen, den Betroffenen (i. d. R. Verbraucher oder Nutzer), der Datenschutzkontrolle sowie der Öffentlichkeit im Rahmen einer demokratischen Überprüfung der Big-Data-Anwendungen.<sup>84</sup>

Die Betroffenen möchten häufig in der Lage sein zu verstehen, warum und welche Big-Data-Verfahren durchgeführt werden und zu welchen Ergebnissen sie führen. Dies basiert einerseits auf einem allgemeinen Informationsinteresse bzw. Auskunftersuchen der Verbraucher, das bei datenschutzrechtlich relevanten Big-Data-Verfahren (mit personenbezogenen Daten) auch rechtlich begründet ist. Andererseits sollen die Verbraucher auch Daten korrigieren können<sup>85</sup>,

<sup>82</sup> Vgl. (Steinbach et al, 2015).

<sup>83</sup> Vgl. dazu auch Abschnitt 7.1.

<sup>84</sup> Vgl. (Weichert 2013).

<sup>85</sup> Berichtigung, Sperrung, Löschung, Widerruf.

wenn die Daten oder Anwendungen fehlerhaft oder ungenau sind oder ggf. zu falschen Schlussfolgerungen führen. Rechtlich durchsetzbar sind diese Interessen der Verbraucher in Bezug auf Transparenz nur bei Big-Data-Verfahren mit personenbezogenen bzw. teilweise auch mit pseudonymisierten Daten.

Big-Data-Verfahren werden aber häufig gerade auch anonym durchgeführt, da diese Anwendungen in bestimmten Bereichen mächtiger sein können. Hier können die Betreiber selbst entscheiden, welche Transparenzanforderungen sie erfüllen möchten, da es keine rechtlichen Anforderungen dafür gibt.

## Dimensionen von Transparenz

Transparenz zu Big-Data-Verfahren umfasst mehrere Dimensionen. Transparenz hat neben der Funktion für die Betroffenen auch eine gesamtgesellschaftlich-demokratische Funktion. Sie muss dann gegenüber der Öffentlichkeit oder zumindest gegenüber unabhängiger Forschung oder Kontrollbehörde vorliegen. Hierbei besteht ein Konflikt mit dem Schutz von Betriebs- und Geschäftsgeheimnissen derjenigen, die Big-Data-Verfahren anwenden. Dieser Konflikt kann durch Marktregulierung oder andere Regulierung gesteuert werden.

Transparenz für den Betroffenen ist bedeutend für die informationelle Selbstbestimmung der Betroffenen. Da bei Big Data i.d.R. sehr viele betroffen sind, wird häufig eine öffentliche Transparenz gefordert, die Unternehmen aber mit ihrem Interesse auf Wahrung der Betriebs- und Geschäftsgeheimnisse abgleichen müssen.

Transparenz kann direkt die Qualität der verwendeten Daten beeinflussen. Nur wenn die einfließenden Daten den Betroffenen bekannt sind, können sie diese auf ihre Aktualität und Richtigkeit hin überprüfen. Dies kann die Richtigkeit der Big-Data-Verfahren beeinflussen.

## Aspekte von Transparenz

Konkret kann Transparenz bei Big-Data-Verfahren unter anderem die folgenden Aspekte umfassen:

- Informationen leicht zugänglich, jederzeit verfügbar für Beteiligte,
- stetige Information, welche Geräte welche Daten sammeln, nutzen und übertragen,
- Informationen über Umfang der harten Daten bzw. eigenen Hochrechnungen/Prognosen,
- Informationen über Algorithmen,
- Aufklärung über Entscheidungen auf Basis der Big-Data-Verfahren bzw. -Algorithmen,
- Informationen über Änderungen,
- Versionierung der Informationen.

## Wirkungen von Transparenz

Sofern Transparenz von den Betreibern der Big-Data-Verfahren geleistet wird, ermöglicht diese Transparenz u. a.

- Widerspruchsmöglichkeiten für Nutzer (die bei personenbezogenen Daten und pseudonymen Nutzerprofilen bei Telemedien gesetzlich gefordert ist),
- die Offenlegung von Quellen und Herkunft der Daten,
- die Einstellung von Profilmerkmalen durch die Nutzer.

Diese Aspekte können die Qualität der Big-Data-Verfahren – z. B. bei sehr vielen Widersprüchen – einschränken. Allerdings können dadurch auch die Big-Data-Verfahren qualitativ verbessert werden, wenn die Nutzer selber Fehler feststellen und korrigieren.

## 7.4 Wie Analyseergebnisse zustande kommen

In immer mehr Organisationen werten eigens dafür abgestellte Experten die Datenbestände systematisch aus. Dieses »Data Scientists« orientieren sich dabei an bewährten Vorgehensmodellen. Am Anfang steht das Verständnis der Business-Anforderungen und ihrer Übersetzung in Fragestellungen für Datenanalyse. Nach der Analyse der verfügbaren Datenquellen startet der Data Scientist die Entwicklung einer datengetriebenen Lösung, die meist in mehrere Iterationsstufen verfeinert wird. Die Lösung wird in einem letzten Schritt in die Prozesse und Applikationen integriert und nachfolgend weiter optimiert.

### Anforderungen an Data Scientists

Die Hauptaufgaben des Data Scientists bestehen darin, Business-Anforderungen in Fragestellungen für Datenanalyse zu übersetzen, diese umzusetzen und zu verfeinern und schließlich den Einsatz der gefundenen Lösungen zu begleiten.

Aus diesem Grund muss der Data Scientist eine Reihe von Fähigkeiten besitzen, die bis vor einigen Jahren noch auf verschiedene Rollen verteilt waren. Er muss sich sowohl auf der Business-Seite, als auch in der Datenanalyse auskennen und zusätzlich das technische Wissen mitbringen, um die Lösungen robust und skalierbar umzusetzen.

Diese Integration wurde nötig, weil sich in den letzten Jahren die technischen Möglichkeiten so schnell entwickelt haben, dass nur durch eine integrierte Sicht auf alle Aspekte gute Lösungen finden lassen. Funktionierende Big-Data-Lösungen kann nur entwickeln, wer weiß, welche Datenanalysemethoden für eine bestimmte Fragestellung existieren und welche technischen Eigenschaften diese Methoden aufweisen.

Data Scientists finden ihr Einsatzgebiet oft in technologiegetriebenen Unternehmen, in denen ständig große Mengen an Daten anfallen, so dass es große Potenziale gibt, durch Auswertung

von Daten wichtige Einsichten zu gewinnen. Dies betrifft insbesondere Firmen, die hauptsächlich über eine Webseite oder ähnliche technischen Services mit ihren Verbrauchern interagieren.<sup>86</sup>

## Phasen der Projektentwicklung

Am Anfang eines Projektes besteht die Aufgabe darin, zusammen mit den entsprechenden Geschäftsbereichen herauszufinden, welche Fragestellungen vorliegen und wie diese mit Hilfe von Datenanalyse gelöst werden können. Hierzu müssen die vorhandenen Datenquellen identifiziert werden, die zusammen mit entsprechenden Datenanalysemethoden zur Problemlösung beitragen können.

Von diesem ersten Ansatz ausgehend, startet der Data Scientist dann die Entwicklung und Verfeinerung einer datengetriebenen Lösung<sup>87</sup>, beginnend mit einer repräsentativen Teilmenge aller bereits verfügbaren Daten. Wichtig ist hierbei, bereits durch Simulation auf vorhandenen Daten die Performance auf zukünftigen Daten robust vorhersagen und abschätzen zu können. Im Laufe dieses Prozesses, der gewöhnlich mehrere Iterationsstufen durchläuft, werden Datenquellen, Datenanalysemethoden und auch die ursprüngliche Fragestellung so lange verfeinert, bis eine Lösung vorliegt, die auf Simulationen und vorhandenen Daten ausreichend gute Ergebnisse liefert.

Als nächste Stufe muss diese Lösung dann in das bestehende System integriert werden, um weitere Evaluierungsergebnisse zu erhalten. Wichtig ist dabei, dass sich die Frage, ob ein System funktioniert, nur auf Basis realer Daten beantworten lässt. Der Data Scientist hat hier also eine besondere Verantwortung, die Evaluierung fachgerecht durchzuführen.

Als letzter Schritt in diesem Prozess besteht dann der Einsatz der gefundenen Lösung im Gesamtsystem, wobei neue Aspekte wie Skalierbarkeit, Monitoring, usw. hinzutreten. Auch jetzt ist das Projekt noch nicht abgeschlossen, sondern bietet langfristig Potenzial für Verbesserung und weitergehende Anpassungen.

## Data Science und Big Data

Da der Data Scientist potenziell mit Daten aus verschiedensten Quellen arbeitet, muss er auch für Datensicherheit und -schutz sensibilisiert sein.

Das Grundinteresse des Data Scientists liegt in der Entwicklung automatisierter Methoden, mit denen sich geschäftsrelevante Fragestellungen skalierbar lösen lassen. Die individuelle Suche nach einzelnen Fragen, zum Beispiel um über einen Nutzer gewissen Antworten zu erhalten, liegt nicht in seinem Fokus.

---

<sup>86</sup> Für einen Webshop kann Data Science zum Beispiel eingesetzt werden, um häufig zusammen erworbene Artikel zu identifizieren, um sie dem Kunden im geeigneten Kontext anzubieten.

<sup>87</sup> Zur Ethik von Algorithmen vgl. (Wagner, 2015).

Zweifellos besitzen moderne Datenanalysemethoden ein großes Problemlösungspotenzial. Es darf jedoch nicht übersehen werden, dass deren »Intelligenz« beschränkt ist. Die Methoden beruhen auf der massiven Anwendung statistischer Methoden und Modelle, die keine inhärente Intelligenz besitzen.<sup>88</sup> Computer bilden sich also weder eine Meinung über die Daten noch »verstehen« sie sie in irgendeiner Form.

Data Science und Big Data sind prinzipiell verschiedene Bereiche. Data Science kann bereits für kleine Datenmengen sinnvoll und anspruchsvoll sein, während Big Data hauptsächlich Technologien umfasst, die in der Lage sind, große Datenmengen skalierbar zu speichern und verarbeiten. Die große Datenmenge macht den Einsatz spezieller, skalierbarer Datenanalysemethoden nötig, und durch den Datenreichtum ergeben sich auch leicht andere Herangehensweisen.

## 7.5 Bedeutung ethischer Grundsätze bei Big-Data-Lösungen

Die Akzeptanz von Big-Data-Technologien in der breiten Öffentlichkeit setzt insbesondere voraus, dass die Technologieanbieter eine ethisch fundierte und untermauerte Herangehensweise praktizieren. Ein verantwortungsvoller Umgang mit Big Data zeichnet sich dadurch aus, dass die Rechte und Interessen der Betroffenen bei der Verwendung von Big Data gewahrt werden und damit ein ethisches Korrektiv zu den nahezu unbegrenzten technischen Möglichkeiten besteht.

### Ethisch-moralische Gesichtspunkte und Transparenz

In Teilen der Gesellschaft und der Politik bestehen Vorbehalte gegenüber Big-Data-Anwendungen.<sup>89</sup> Die Medien rücken unter dem Schlagwort Big Data teilweise Gefahren und negative Beispielfälle in den Vordergrund (»Big Brother«). Die Verbraucher verlangen, berechtigterweise,

88 Zum Beispiel kommen für die Verhaltensvorhersage eines Nutzers Modelle zum Einsatz, die im Wesentlichen einfach strukturierte Funktionsmodelle mit einer großen Anzahl an Parametern sind, die im Laufe des Trainingsprozesses lediglich so angepasst werden, dass die Vorhersage auf den verfügbaren Trainingsdaten möglichst gut ist.

89 Einige Autoren richten ihre Aufmerksamkeit auf grundsätzliche Herausforderungen im Zusammenhang mit Big Data. Sie befürchten, dass die Datafizierung Chancen für eine alternative Zukunft verbaut. So geht es Morozov (vgl. (Morozov, 2015)) u. a. um ein neues Verständnis von Datenschutz, die »wirmächtige Verschmelzung kommerzieller Interessen heutiger Unternehmen mit den Sicherheitsinteressen heutiger Regierungen« (vgl. (Morozov, 2015, S. 4)) als größte Herausforderung sowie die Frage, »ob es angemessen ist, dass es die großen Unternehmen wie Google und Facebook sind, die letztgültig vermitteln, wie wir leben, heilen, studieren oder reisen.« (vgl. (Morozov, 2015, S. 7)) Zum Datenschutz lautet Morozovs zentrale Aussage. »Sollte es gelingen, uns von diesem theoretischen Vorurteil freizumachen, dass es beim Datenschutz um den Schutz eines Gutes aus der Vergangenheit geht, und würden wir Datenschutz stattdessen als Möglichkeit betrachten, eine alternative Zukunft zu leben – eine, die uns von niemandem aufgezwungen wird, sondern die wir uns selbst teilautonom wählen – dann würden wir erkennen, dass die gegenwärtigen Prozesse, die alles quantifizieren und datafizieren, diese Möglichkeit stark einschränken.« (vgl. (Morozov, 2015, S. 3f.)) Ähnlich argumentiert Hofstetter: »Die Macht verschiebt sich weg vom demokratisch legitimierten Staat, dessen Repräsentanten durch den Souverän wählbar und kontrollierbar sind, hin zu den Wirtschaftsbetrieben, die über unsere Daten verfügen – und die Schlüsseltechnologien, um sie auszuwerten. Anders als ein demokratischer Staat sind sie hochgradig intransparent; durch ihre Geschäftsstrategie, stets ein Monopol anzustreben, wird Nutzern jede Alternative genommen.« (Vgl. (Hofstetter, 2015, S. 34))Vgl. (Bitkom 2012, S. 14, 16).



nach einer transparenten Verarbeitung ihrer Daten und lassen den pauschalen Hinweis auf Geschäftsgeheimnisse nicht gelten. Schließlich stellen politische Entscheidungsträger laufend wechselnde Forderungen an die Wirtschaft. Auch ein Großteil der im Zusammenhang mit Big Data entstehenden Rechtsfragen ist noch nicht abschließend geklärt, so dass auch hieraus Unsicherheit entsteht. Schließlich müssen aber vor allem die Unternehmen selbst sicherstellen, dass sie eigenen Wertvorstellungen treu bleiben und entsprechend ihre Geschäftsprozesse gestalten (vgl. Leitlinie 7).

Der Grundsatz bei Big Data lautet also: Nicht alles, was machbar ist, wird auch umgesetzt. Mit anderen Worten: Auch wenn eine Big-Data-Lösung sowohl technisch als auch juristisch umgesetzt werden könnte, so müssen dennoch beispielsweise ethisch-moralische Gesichtspunkte bei der Letztentscheidung berücksichtigt werden (vgl. Leitlinie 9).

Vor diesem Hintergrund ist es für Unternehmen besonders wichtig, gegenüber ihren Verbrauchern transparent über Big-Data-Lösungen zu informieren. Dies kann über separate Informationen zu Big-Data-Lösungen geschehen oder im Rahmen der geschäftspolitischen Außendarstellung wie Internetauftritt. Transparenz ist ein wesentlicher Schlüssel für die Akzeptanz von Big-Data-Lösungen in der öffentlichen Meinungsbildung (vgl. Leitlinie 2).

## Respekt vor Verbrauchern

Eng verzahnt mit der Transparenz ist der Respekt vor Verbrauchern. In der Regel setzen Big-Data-Lösungen auf Daten auf, welche das Unternehmen über die eigenen Kunden aufgrund einer Vertragsbeziehung hat.<sup>90</sup> Um die Akzeptanz der Kunden für eine solche Datenverarbeitung zu erhalten, muss dem Kunden nachvollziehbar erläutert werden, wie und wofür die Daten verwendet werden. Letztendlich partizipieren beide Seiten am Nutzen von Big-Data-Lösungen (vgl. Leitlinie 1) – zum einen das Unternehmen und zum anderen der Kunde. Das Unternehmen kann den wirtschaftlichen Wert durch Erstellung von innovativen Produkten erhöhen. Gleichzeitig profitiert der Verbraucher von optimierten und nützlicheren Produkten zu einem besseren Preis.

## Ethisch-moralisches Korrektiv

IT-Unternehmen befinden sich in einem größer werdenden Spannungsfeld. Um die sich bietenden wirtschaftlichen Chancen ausnutzen zu können, müssen sie den vielfältiger werdenden Möglichkeiten des Einsatzes von Big Data aufgeschlossen gegenüber treten. Andererseits wird sich nachhaltiger Erfolg mit Big-Data-Technologien nur einstellen, wenn die Unternehmen hiermit auf Akzeptanz bei den Nutzern stoßen. Eine zentrale Voraussetzung dafür ist eine ethisch fundierte und untermauerte Herangehensweise.

Auch wenn die Optimierung von Produkten als wirtschaftliches Ziel für Unternehmen sehr erstrebenswert ist, so muss bei der Kundenentwicklung gerade bei Big-Data-Lösungen ein

---

<sup>90</sup> Vgl. auch (Ulmer, 2013, S. 225). Zu allgemeinen Methoden der Anonymisierung vgl. (Dammann 2014).

ethisch-moralisches Korrektiv Berücksichtigung finden. Big Data ermöglicht per definitionem die Verknüpfung und Analyse unterschiedlichster Daten aus einer Vielzahl von Datenquellen. Eine solche maximale Verknüpfung, Datenschutzkonformität unterstellt, kann zu unerwarteten Produktentwicklungen führen. Die Verknüpfung unterschiedlichster Datenquellen könnte Ergebnisse hervorbringen, die ethisch-moralisch bedenklich sind. Besondere Sorgfalt ist geboten, wenn »besondere Daten«, gemäß § 3 Abs. 9 BDSG<sup>91</sup>, verarbeitet und analysiert werden. Daher ist vor der Verwendung solcher Daten genau zu prüfen, ob das angestrebte Ergebnis auch unter ethischen Gesichtspunkten vertretbar bzw. erwünscht und im Sinne derer ist, von denen die Daten stammen. Wenn nicht klar ist, welche Konsequenzen sich aus einer Verwendung oder Verknüpfung dieser Daten ergeben können, sollten Unternehmen von der Verwendung absehen. (vgl. Leitlinie 9).

## Bedeutung von Leitlinien

Big-Data-Analysen bergen also vielfältige Herausforderungen, nicht zuletzt technischer und wirtschaftlicher Natur.<sup>92</sup> Die im Kapitel 8 vorgestellten Leitlinien sollen Hinweise für die Kommunikation mit Verbrauchern geben, Best Practices für das Prozessmanagement und die Unternehmensorganisation bieten und den Unternehmen insgesamt eine Richtschnur für einen ethisch und juristisch einwandfreien Einsatz von Big-Data-Analysen aufzeigen.

Sie können für die Bewältigung nur einiger dieser Herausforderungen Hinweise geben. Vorge stellt werden insbesondere Möglichkeiten, einen verantwortungsvollen Umgang mit Big Data zu praktizieren, der sich dadurch auszeichnet, dass die Rechte und Interessen der Betroffenen bei der Verwendung von Big Data – sei es in den eigenen Geschäftsprozessen oder bei der Gestaltung von Produkten – angemessen berücksichtigt werden und damit ein ethisches Korrektiv zu den nahezu unbegrenzten technischen Möglichkeiten besteht (vgl. Leitlinie 7).

## 7.6 Stimulierung gesellschaftlich wünschenswerter Verhaltensweisen

Individuelle risikoorientierte Versicherungstarife stehen dem deutschen Solidarprinzip nicht zwingend im Weg. Sie eröffnen die Möglichkeit, ein gewünschtes Verhalten zu fördern, ohne die Risiken völlig auf das Individuum abzuwälzen. Big-Data-Analysen können auch einen digitalisierten Schutz ermöglichen, indem auf ihrer Grundlage Empfehlungen generiert und digitalisiert übertragen werden. Einzelne Verbraucher sollten zwischen der Ablehnung einer individuellen Vermessung ihrer Gewohnheiten bis hin zur Akzeptanz und freiwilligen Datenlieferung wählen können.

91 »Besondere Arten personenbezogener Daten sind Angaben über die rassische und ethnische Herkunft, politische Meinungen, religiöse oder philosophische Überzeugungen, Gewerkschaftszugehörigkeit, Gesundheit oder Sexualleben.«

92 Die technischen und wirtschaftlichen Herausforderungen sind nicht Gegenstand der Leitlinien.

## Stimulierung – ja, Sanktionierung – nein

Big-Data-Anwendungen sollten so konzipiert werden, dass sie ggf. eine Stimulierung, nicht jedoch eine Sanktionierung ermöglichen. Dabei muss dafür Sorge getragen werden, dass Personen, die in eine Verarbeitung ihrer Daten nicht einwilligen, keine Nachteile haben.

Zu klären ist, unter welchen Umständen eine Benachteiligung und eine Sanktionierung durch eine Anwendung vorliegen.

Personalisierte, risikoorientierte Versicherungstarife müssen dem deutschen Solidaritätsprinzip im Versicherungswesen nicht zwingend widersprechen, wenn die Ergebnisse der Datenanalyse lediglich zur Stimulierung dienen. An die Stelle von Negativkonsequenzen können Empfehlungen an die Verbraucher rücken, wie sie den definierten Kriterien besser entsprechen.<sup>93</sup> Big Data kann hier im Sinne des Nudging<sup>94</sup> wirken; dabei muss ein ausgewogenes Verhältnis zwischen Selbstverantwortung und Solidargemeinschaft bestehen bleiben.

Sollte bei einer Anwendung Unsicherheit herrschen, ob sie stimuliert oder sanktioniert, könnte eine Ethik-Kommission zu Rate gezogen werden. Unterstützung bei dieser Entscheidung sollen auch die Leitlinien (vgl. Kapitel 8) bieten.

## 7.7 Datensammlung im Automatismus

Die automatisierte Sammlung von Daten bietet sowohl für den Verbraucher als auch für Unternehmen einen erheblichen Mehrwert. Zu gewährleisten ist dabei aber, dass die rechtlichen Vorgaben beachtet werden und der Verbraucher darüber entscheiden kann, welche Daten seine Geräte übermitteln sollen.

### Automatisierte Sammlung von Daten

Die automatisierte Sammlung von Daten birgt einen erheblichen gesellschaftlichen Mehrwert. Mit der Ausbreitung des Internets der Dinge vervielfachen sich die Möglichkeiten zu einer solchen Datensammlung über Sensoren, die in Alltagsgegenständen eingebaut sind, seien es Fitnessarmbänder, Turnschuhe oder Waschmaschinen. Entscheidend ist dabei jedoch, dass Unternehmen die Interessen der Betroffenen angemessen berücksichtigen. Gerade die ubiquitäre automatisierte Datensammlung erfordert ein ethisch einwandfreies Vorgehen, weil sie eine stärkere technische Durchdringung des Bereichs privater Lebensgestaltung mit sich bringt (vgl. Leitlinie 7 und Leitlinie 8).

<sup>93</sup> Eine solche Empfehlung könnte beispielsweise zur Teilnahme an einem Rückentrainingszirkel oder Fahr-sicherheitstraining sowie zur Wahl einer anderen, altersgerechteren Wanderroute oder einer alternativen Fahrstrecke, aufgrund erhöhten Unfallvorkommens anleiten.

<sup>94</sup> Nachhaltige Verhaltensbeeinflussung durch Impulse (wörtlich: Schubs) – ohne Verbote oder Befehle.

## Inhaberschaft an Daten

Dabei sind insbesondere die Rechte der Verbraucher zu beachten. Data Mining mit personenbezogenen Daten darf nicht ohne die Kenntnis der Betroffenen erfolgen. Um die Akzeptanz der Verbraucher für die Datenverarbeitung zu sichern, muss diesen gegenüber transparent gehandelt werden. Der Verbraucher kann sich nur dann selbstbestimmt und ohne ein »diffuses Gefühl des Beobachtetwerdens« den neuen Technologien öffnen. Insbesondere ist auch zu bedenken, dass die Inhaberschaft an Daten rechtlich nicht geklärt ist. Die technische Umsetzbarkeit der Datenerhebung begründet für sich noch keine ethische Legitimität. Vielmehr müssen sowohl das Daten-subjekt als auch der Eigentümer des Geräts, welches die Daten sammelt, über die Datenerhebung informiert sein. Sie müssen die Kontrolle über die ihnen vom Gesetz zugeordneten Herrschaftssphären behalten.

## Technische Ausgestaltung der Datensammlung

Besondere Bedeutung kommt hierbei schon der technischen Ausgestaltung der Datensammlung zu. Die beiden wichtigsten Aspekte lassen sich unter den Schlagworten »Privacy by Design« und »Privacy by Default« zusammenfassen. Schon bei der technischen Konstruktion informationstechnischer Systeme, die der Verarbeitung von personenbezogenen Daten dienen, muss sichergestellt sein, dass das System datenschutzfreundlich eingesetzt werden kann. Nicht nur stellt das Recht auf informationelle Selbstbestimmung ein Grundrecht dar, verfassungsgerichtlich sind mittlerweile auch Anforderungen an die Ausgestaltung informationstechnischer Systeme festgeschrieben. Soweit möglich sollten Konflikte mit diesbezüglichen rechtlichen Vorgaben deshalb schon durch technische Gestaltung minimiert werden. Darüber hinaus sollten Unternehmen, gerade bei sensiblen Daten, nicht standardmäßig die maximal mögliche Datenmenge erheben, sondern eine datenschutzfreundliche Voreinstellung wählen.

Der Entwurf der Datenschutz-Grundverordnung enthält die Vorgabe, »Privacy by Design« und »Privacy by Default« einzusetzen. Weitere Ansätze, konkrete Aspekte von »Privacy by Design« und »Privacy by Default« regulatorisch zu verankern, finden sich in den gesetzlichen Vorgaben zum Einsatz von intelligenten Messsystemen (Smart Meter). Die entsprechenden Regelungen sehen derartige Schutzmechanismen als verbindlich vor. Sie können eine Orientierung auch außerhalb der Energiewirtschaft bieten.

## 7.8 Recht auf Vergessenwerden?

Die Forderung eines »Rechts auf Vergessenwerden« geht an der technologischen Realität vorbei. Vielmehr geht es um selbstbestimmtes Löschen und darum, überhaupt zu erfahren, wo welche personenbezogenen Daten liegen.

Für die Speicherung von massenhaft anfallenden Daten bilden die Kosten mittlerweile keine Barriere mehr – mit der Folge, dass Daten für die Ewigkeit vorgehalten werden können. Das

Internet vergisst nicht: Digital veröffentlichte Bilder oder Meinungsäußerungen – auf Unternehmensservern gespeichert und im Internet veröffentlicht – verschwinden von dort nicht mehr ohne aktives Handeln. In der Praxis kann das für Personen die Konsequenz haben, dass z. B. ihre politischen Äußerungen ihnen über Jahrzehnte anhängen, selbst wenn sie ihren Überzeugungen längst abgeschworen haben. Daraus können erhebliche Nachteile<sup>95</sup> erwachsen. Diese Möglichkeit kann das Verhalten in der Gegenwart deutlich beeinflussen, um in Zukunft nicht »bestraft« zu werden. Die Löschung von Inhalten im Internet bleibt eine Herausforderung, wenn man nicht alle Verbreitungswege kennt.<sup>96</sup>

## Verfallsdatum für Informationen?

Im Sinne der informationellen Selbstbestimmung wird in der Öffentlichkeit die Forderung laut,

- die Löschung von personenbezogenen Daten veranlassen zu können bzw.
- dass Daten nur für einen bestimmten Zeitraum gespeichert werden.

Digitale Informationen mit Personenbezug sollen nicht dauerhaft verfügbar sein. Ein »Recht auf Vergessenwerden« postulierte Viktor Mayer-Schönberger.<sup>97</sup> Er stellte die Forderung auf, dass elektronisch gespeicherte Daten mit einem Verfallsdatum versehen und nach dessen Ablauf gelöscht werden. Mayer-Schönberger geht dabei nicht zwingend von einem Automatismus oder einem tatsächlichen Vergessensrecht aus, ebenso wenig von einem umfassenden Data Rights Management System. Vielmehr forderte er die permanente Konfrontation mit der Frage, wie lange ein Datum gespeichert werden soll. Schrittweise würde das Bewusstsein dafür geschärft, dass Daten und enthaltene Informationen einen zeitlich begrenzten Wert aufweisen und folglich manches nicht ewig gespeichert werden muss. Digitales Merken und Vergessen soll funktionieren wie analoges. Analog ist es eine bewusste Entscheidung und eher die Ausnahme, etwas im Geiste oder extern festzuhalten, wenn es wichtig erscheint. Das Vergessen ist hier die Regel und das Merken die Ausnahme. Digital verhält es sich umgekehrt: Das Merken funktioniert automatisiert, wohingegen Vergessen einer aktiven Handlung bedarf.<sup>98</sup>

## Unklare Rechtslage

Gesetzlich ist die Forderung nach einem Recht auf Vergessenwerden bis dato nicht geregelt. In der geltenden europäischen Richtlinie und im deutschen Datenschutzgesetz werden nur Voraussetzungen bestimmt, unter denen ein personenbezogenes Datum zu löschen ist. 2011 wurde das Recht im Vorschlag der Europäischen Kommission für eine Datenschutz-Grundverordnung

---

95 Nachteile für Individuen könnten beispielsweise in Form von schlechteren Kreditkonditionen bei Banken oder von schlechteren Tarifen bei Versicherungen oder in Form einer Überwachung durch staatliche Institutionen entstehen. Die Interessenabwägung ist alles andere als trivial, verfügen doch Banken und Versicherungen über valide Grundlagen für eine unterschiedliche Bewertung von Verbrauchern.

96 Wird ein Verweis an der einen Stelle gelöscht, kann er an einer anderen nach wie vor bestehen. Selbst wenn ein alkoholisierte junger Partyteilnehmer leichtsinnig veröffentlichte Bilder bereits nach kurzer Zeit in seinem Profil löscht, können diese Bilder Jahre später wieder auftauchen und einen Karriereknick verursachen.

97 Vgl. (Mayer-Schönberger, 2010).

98 Vgl. (Pluta, 2008).

aufgegriffen, wobei keine Löschrufen vorgesehen wurden und das Recht vornehmlich einer Ausweitung des Rechts auf informationelle Selbstbestimmung gleichkam. Aufgrund zahlreicher Kritik wurde das Recht 2013 aus dem Entwurf entfernt und auf ein Recht auf Löschung beschränkt.<sup>99</sup>

Grundsätzlich besagt das geltende Recht, dass Daten nicht länger personenbezogen aufbewahrt werden dürfen, als es für die Realisierung der Zwecke erforderlich ist, für die sie erhoben wurden oder weiterverarbeitet werden. Diese Formulierung lässt Spielraum für die Ausgestaltung von unternehmensspezifischen Datenschutzrichtlinien und -erklärungen. Daher gibt es unterschiedliche Handhabungen in den Unternehmen, die nicht unbedingt zu mehr Transparenz und Kontrolle für den Verbraucher führen. Jedoch starteten Unternehmen in den letzten Jahren Initiativen für die Entwicklung standardisierter Löschkonzepte, die sicherstellen sollen, dass jedes Datum gelöscht wird, sobald es nicht mehr erforderlich ist.<sup>100</sup>

2014 entschied der EuGH, dass es für den Einzelnen möglich sein muss, die Löschung von Links auf personenbezogene Daten zu veranlassen<sup>101</sup>. Diese Entscheidung wurde so umgesetzt, dass die betreffenden URLs aus den Suchergebnissen ausgeblendet werden. Die tatsächlichen verlinkten Inhalte liegen nach wie vor auf den Unternehmensservern. Archive von Medien sind folglich nicht betroffen. Die genannte Limitierung haben alle bisherigen technischen Ansätze gemeinsam. Zwar geht aus der öffentlichen Debatte hervor, dass dies oft als positiv empfunden wird, da das Recht auf Vergessenwerden in seiner umfassenden Form eine Zensur darstellen und massive Erinnerungslücken erzeugen könne.<sup>102</sup> Jedoch stellt die Limitierung trotzdem eine massive Einschränkung der informationellen Selbstbestimmung dar. Umgekehrt würde ein absolutes Löschrufen einen entsprechenden Eingriff in die Informations- und Pressefreiheit darstellen. Wie so oft in dieser Diskussion hat man es mit widerstreitenden berechtigten Interessen und Grundrechten zu tun.

## Löschung von Daten bleibt eine Herausforderung

Wegen der kontroversen Interessen von Medien, Unternehmen und den Betroffenen erscheint die Forderung nach einem Recht auf durchgreifendes Löschrufen utopisch. Vielmehr ist Kompromiss zu finden. Vor diesem Hintergrund erscheint Mayer-Schönbergers Ansatz sinnvoll, da er nicht von einer rechtlichen Festlegung ausgeht, sondern von einem Bewusstseinswandel. In der öffentlichen Debatte geht es auch meist viel weniger um das Einklagen eines Rechts, als vielmehr um Möglichkeit zu selbstbestimmtem Löschrufen verfügen. Hierzu müssen Strukturen geschaffen werden, die die Verbreitung von Daten im Internet nachvollziehbar machen. Grundsätzlich spielt hierbei auch die Forderung nach erhöhter Transparenz über die Speicherung und Verwendung personenbezogener Daten eine Rolle. So wird erörtert, ob Betroffene als eine Art

---

99 Auf technischer Ebene existieren erste, allerdings noch nicht ausgereifte Ansätze, die die Vorschläge Mayer-Schönbergers verfolgen. Vgl. (Heise online, 2011) sowie (Welchering, 2011).

100 Vgl. (DIN 2015).

101 Vgl. (EuGH, 2014).

102 Vgl. (Reißmann, 2014).

»Datenbestätigung« in bestimmten Abständen darüber informiert werden könnten, welche Daten gespeichert, genutzt und verarbeitet werden und zu welchem Zweck das geschieht. Dabei müsste auch die Weitergabe an Dritte weitgehend transparent gemacht werden, so dass die Datenverbreitung nachvollzogen werden kann. Ein solches Vorgehen würde gewährleisten, dass Betroffene notwendige Korrekturen an veralteten oder falschen Beständen aufzeigen könnten. Allerdings wäre eine aktive Benachrichtigung aller Betroffenen in regelmäßigen Abständen mit großen praktischen und auch datenschutzrechtlichen Schwierigkeiten verbunden. Es gibt jedoch recht erfolgreiche Ansätze, bei denen im Rahmen eines Nutzeraccounts jederzeit für den Nutzer einsehbar ist, welche seiner Daten gespeichert sind und wo er auch selbst deren Berichtigung oder Löschung initiieren kann. In dieser Richtung können sicherlich noch weitere Instrumente entwickelt werden, mit denen die notwendige Transparenz und Interaktion hergestellt werden kann, die für die Nutzer zur Ausübung ihres informationellen Selbstbestimmungsrechts notwendig ist.

## 7.9 Umgang mit Wahrscheinlichkeiten und Verhaltensprognosen

Grundsätzlich ist es zu begrüßen, wenn mit Hilfe von Big Data Prognosen verbessert werden. Prognosen können jedoch an Grenzen geraten: Sie dürfen Menschen nicht ihre Chancengleichheit nehmen.

Big Data Analytics macht es möglich, aus bekannten Umständen Prognosen über das künftige Verhalten einer Person oder andere Ereignisse abzugeben. Das gilt für die Rückzahlung von Darlehen genauso wie die Wahrscheinlichkeit eines Autounfalls.

Grundsätzlich ist es zu begrüßen, wenn mit Hilfe von Big Data Prognosen verbessert werden. Die Verbesserung einer Prognose über die Bonität eines Verbrauchers ist genauso im Interesse der Bank, die das Geld zurückerhalten will, wie im Interesse des Verbrauchers, der am besten nur so viel Schulden aufnimmt, wie er wahrscheinlich auch zurückzahlen kann. Sie ist auch im Interesse anderer Verbraucher, denn je besser die Prognose, desto geringer können die Risikoaufschläge sein, die Banken erheben müssen.

Auch für Gesundheitsprognosen gilt das. Je besser man voraussehen kann, welche Krankheiten ein Mensch möglicherweise entwickelt, desto früher kann man therapeutische Gegenmaßnahmen ergreifen. Das nützt vor allem den Betroffenen, es nützt aber auch der Gesellschaft durch eingesparte Kosten, die für andere Maßnahmen ausgegeben werden können.

### Grenzen von Prognosen

Prognosen können aber auch ungewollte Folgen haben. Bezieht man in eine Bonitätsprüfung auch den Wohnort des Verbrauchers mit ein, so kann dies bedeuten, dass Menschen aus bestimmten Stadtgebieten keine Kredite mehr erhalten, vielleicht nicht einmal einen Mobilfunkvertrag. Prognosen könnten auch dazu führen, dass Menschen mit einer bestimmten Erbanlage keine Krankenversicherung abschließen können, obwohl sie vielleicht diese Krankheit

niemals entwickeln; so würde das Solidaritätsprinzip von Versicherungen ausgehebelt. Hier geraten Prognosen an ihre Grenzen. Prognosen dürfen nicht dazu führen, dass Menschen die Chancengleichheit genommen wird (vgl. Leitlinie 9). Bestimmte Analysen müssen ausgeschlossen bleiben, beispielsweise genetische Analysen für berufliche Zwecke. Auch dürfen staatliche Zwangsmaßnahmen nicht mit einer bloßen Wahrscheinlichkeitsprognose begründet werden: Die Polizei darf einen Fußballfan nicht deshalb festnehmen, weil die Prognose eine hohe Wahrscheinlichkeit ergab, dass er im Stadion randaliert. Staatliches Handeln darf nur an begangene Handlungen anknüpfen oder an eine bestehende Gefahr, die durch konkrete Tatsachen belegt werden muss.

## 7.10 Erkenntnisse aus der Kommunikation in Sozialen Netzwerken

Unternehmen sollten bei der Analyse der Kommunikation in Sozialen Netzwerken umsichtig agieren und die Interessen der Betroffenen angemessen berücksichtigen.

Obwohl die Kommunikation in Sozialen Netzwerken öffentlich verfügbar ist, sollten Unternehmen bei der Sammlung und Auswertung solcher Informationen berücksichtigen, dass die Betroffenen bewusst meist nur mit einem eingeschränkten Nutzerkreis und nur in einem ganz spezifischen Zusammenhang kommuniziert haben. Nicht immer mag ein Betroffener davon ausgegangen sein, dass seine Äußerungen und Aktivitäten Gegenstand von kommerziellen Auswertungen auch außerhalb des jeweiligen Forums werden, auch wenn die Informationen allgemein zugänglich sind. Deshalb sollten Unternehmen hier umsichtig agieren und die Interessen der Betroffenen angemessen berücksichtigen (vgl. Leitlinie 7). Selbstverständlich sind unbedingt die rechtlichen Rahmenbedingungen zu beachten. Vorzugswürdig ist stets die Analyse nur anonymisierter Daten, bei denen die einzelne Person nicht mehr identifiziert werden kann (vgl. Leitlinie 3). Oftmals geht die Anonymisierung der Daten dabei nicht mit einem Erkenntnisverlust einher.

Weniger problematisch ist aber die Auswertung von Kommunikation und Aktivitäten der Nutzer, die sich unmittelbar auf das Unternehmen selbst beziehen, beispielsweise weil der Nutzer eine Fanseite besucht oder Nachrichten direkt an das Unternehmen geschickt hat. Letztlich sollte bei jeder Information genau abgewogen werden, welche Interessen das Unternehmen überhaupt hat und welche Interessen der Betroffenen der konkreten Verwendung möglicherweise entgegenstehen.







# 8 Leitlinien für den Big-Data-Einsatz

## 8 Leitlinien für den Big-Data-Einsatz

### **Leitlinie 1 – Nutzen der Big-Data-Anwendungen prüfen**

Big-Data-Anwendungen sollen einen klar erkennbaren Nutzen für die Verbraucher, Kunden oder die Gesellschaft haben.

### **Leitlinie 2 – Anwendungen transparent gestalten**

Big-Data-Anwendungen sollen transparent sein, so dass die Betroffenen erkennen können, welche ihrer personenbezogenen Daten in welcher Weise verarbeitet werden.

### **Leitlinie 3 – Bevorzugt anonymisierte oder pseudonymisierte Daten verarbeiten**

Soweit die Verarbeitung von anonymisierten oder pseudonymisierten Daten denselben Nutzen für die Beteiligten hat, sind solche Verfahren vorzuziehen. Es gibt aber auch Verfahren, die sich nur sinnvoll einsetzen lassen, wenn personenbezogene Daten verwendet werden.

### **Leitlinie 4 – Interessen der Beteiligten abwägen**

Personenbezogene Daten dürfen verarbeitet werden, wenn berechtigte Interessen der verantwortlichen Stelle dies rechtfertigen und keine überwiegenden Interessen der Betroffenen entgegenstehen. Unter denselben Voraussetzungen ist es auch zulässig, Daten zu verwenden, die ursprünglich für einen anderen Zweck erhoben wurden. Liegen diese Voraussetzungen nicht vor, dürfen die Daten nur verarbeitet werden, wenn die Betroffenen einwilligen.

### **Leitlinie 5 – Einwilligungen transparent gestalten**

Wenn die Datenverarbeitung in Big-Data-Verfahren auf eine Einwilligung gestützt wird, muss die Einwilligung transparent sein, damit der Betroffene erkennen kann, welche Daten für welche Zwecke verwendet werden.

### **Leitlinie 6 – Nutzen für Betroffene schaffen**

Big-Data-Anwendungen sollten auch einen Nutzen für Betroffene haben, die ihre Daten für die Bearbeitung zur Verfügung stellen.

### **Leitlinie 7 – Governance für personenbezogene Daten etablieren**

Unternehmen sollten eine starke Governance etablieren, die eine gründliche Überprüfung von Zulässigkeit und Notwendigkeit von Big-Data-Anwendungen garantiert, den verantwortungsvollen Umgang mit Big Data sichert und die Rechte und Interessen der Betroffenen wahrt. Hierbei kommt dem betrieblichen Datenschutzbeauftragten eine wichtige Rolle zu.

### **Leitlinie 8 – Daten wirksam gegen unberechtigte Zugriffe schützen**

Unternehmen, die Big-Data-Anwendungen einsetzen, setzen ausreichende technische und organisatorische Schutzmaßnahmen ein, um unberechtigte Zugriffe auf personenbezogene Daten zu verhindern.

### **Leitlinie 9 – Keine Daten zu ethisch-moralisch unlauteren Zwecken verarbeiten**

Datenerhebungen, Verknüpfung von Daten oder andere Datenverarbeitungen zu ethisch-moralisch unlauteren Zwecken sind zu unterlassen. Das gleiche gilt, wenn die Erhebung, Verknüpfung oder Verarbeitung der Daten den Betroffenen schaden können.

#### **Leitlinie 10 – Datenweitergabe nach Interessenabwägung ermöglichen**

Die Weitergabe von personenbezogenen Daten an Dritte ist mit Einverständnis möglich. Möglich ist sie auch nach einer Interessenabwägung, wobei der Weitergebende die Risiken zu berücksichtigen hat, die sich aus der Zusammenführung mit anderen Datenbeständen beim Empfänger ergeben könnten. Dabei ist sicherzustellen, dass der Betroffene informiert wird.

#### **Leitlinie 11 – Selbstbestimmtes Handeln ermöglichen**

Unternehmen, die Big-Data-Anwendungen einsetzen, ermöglichen dem Betroffenen durch transparente Information über die Anwendungen und durch ergänzende Auskünfte selbstbestimmtes Handeln.

#### **Leitlinie 12 – Politische Rahmenbedingungen vervollkommen – Datenschutz und Daten nutzen neu abwägen**

Big-Data-Anwendungen können einen hohen Nutzen für die Gesellschaft und für jeden Einzelnen haben. Nicht nur Unternehmen sind daher gefordert, sondern auch die Politik. Sie entwickelt die Rahmenbedingungen so weiter, dass Big-Data-Anwendungen in Deutschland und der Europäischen Union sinnvoll eingesetzt werden können, Rechte der Betroffenen angemessen geschützt und ungerechtfertigte regulatorische Hindernisse abgebaut werden.

Die Weiterentwicklung einer datenbasierenden Wirtschaft in Deutschland und der Europäischen Union ist ein wichtiges politisches Ziel. Deutsche Unternehmen dürfen hierbei keinen Wettbewerbsnachteilen gegenüber Unternehmen aus anderen EU-Staaten oder anderen Staaten der Welt ausgesetzt sein.

Die Politik setzt einen Prozess auf, der die Weiterentwicklung des Datenschutzrechts auch nach Verabschiedung der EU-Datenschutz-Grundverordnung zum Ziel hat.

### **Einhaltung der Leitlinien**

Im Dialog zwischen Wirtschaft, Wissenschaft und Politik sollte erörtert werden, ob eine Selbstverpflichtung, ein Kodex oder eine weitere Regulierung zielführend ist.<sup>103</sup>

Es reicht nicht aus, Leitlinien für den Big-Data-Einsatz vorzugeben. Vielmehr muss deren Einhaltung glaubwürdig und wirksam überprüft werden. Dafür stehen zwei Möglichkeiten zur Verfügung:

- die freiwillige Selbstkontrolle sowie
- die Zertifizierung.

Beide Möglichkeiten können auch kombiniert werden.

---

<sup>103</sup> Vgl. (Eckert et al, 2014, S.18).

Die freiwillige Selbstkontrolle ist insbesondere dann geeignet, wenn Rechtsgüter von Verbrauchern betroffen sind, da sich die Verbraucher über eine zentrale Stelle informieren und ggf. Beschwerde einlegen können. Die freiwillige Selbstkontrolle gewährleistet Kontinuität bei der Einhaltung der Leitlinien. Das ist insbesondere von Vorteil, wenn die dauerhafte Einhaltung von organisatorischen Maßnahmen und Verhaltenspflichten sichergestellt werden soll. Die freiwillige Selbstkontrolle gilt als effizient, da die Einhaltung einzelner Punkte der Leitlinien anlassbezogen überprüft werden kann und nicht umfassend im Vorfeld durchgeführt werden muss.

Ist ein hoher und anerkannter Grad an Verbindlichkeit gewünscht bzw. erforderlich, lassen sich aus nationalen und internationalen Standards, insbesondere den »EU-Principles for Better Self and Co-Regulation«<sup>104</sup>, die folgenden Mindestanforderungen bezüglich der Durchsetzung ableiten:

- Verbindliche Verpflichtungserklärungen der teilnehmenden Organisationen,
- Monitoring, Evaluierung und kontinuierliche Weiterentwicklung,
- effektive Beschwerdemechanismen und Streitbeilegung,
- wirksame Sanktionen.

Im Gegensatz zur freiwilligen Selbstkontrolle eignet sich eine Zertifizierung insbesondere dann, wenn es um die Einhaltung von Vorgaben zu einem gewissen Zeitpunkt geht. Die Vorlage eines Zertifikats (Siegel) kann einen Beitrag leisten, öffentliches Vertrauen in den Big-Data-Einsatz aufzubauen.

---

104 Vgl. (European Commission, ohne Jahresangabe).







# 9 Big Data und Datenschutz: Politischer Handlungsbedarf

## 9 Big Data und Datenschutz: Politischer Handlungsbedarf

Bitkom plädiert für einen fortwährenden gesellschaftlichen und politischen Diskurs, um gesellschaftlich akzeptierte Standards zu Big Data zu entwickeln. Chancen und Risiken müssen in diesem Diskurs in einen angemessenen Ausgleich gebracht werden und an diesem Maßstab sollten sich die rechtlichen Vorgaben orientieren. Bisher wurde Big Data lediglich unter Schutzaspekten gesetzlich erfasst. Es gilt jedoch, positive Nutzungsregeln ins Auge zu fassen, die der zentralen Bedeutung von Daten in einer Data Driven Economy gerecht werden (vgl. Leitlinie 12).

Der Entwurf der EU-Datenschutz-Grundverordnung bleibt hier noch hinter den Möglichkeiten und Erfordernissen der Digitalisierung zurück. Um für Industrie 4.0, Big Data, Cognitive Computing und andere Innovationen günstige Rahmenbedingungen zu schaffen, muss ein politischer Prozess für die Weiterentwicklung und praktische Umsetzung der Verordnung aufgesetzt werden.

### Veraltetes Datenschutzrecht – ökonomische Risiken

Die datenschutzrechtlichen Rahmenbedingungen haben mit dem technologischen Fortschritt nicht Schritt gehalten. Das deutsche Datenschutzrecht stammt im Kern aus den 70er Jahren des 20. Jahrhunderts. Big Data war damals noch Science-Fiction. Zwar können die abstrakten rechtlichen Regelungen durchaus auch auf neue Technologien angewendet werden, das geht aber mit erheblichen Rechtsunsicherheiten einher, die noch dadurch verstärkt werden, dass es in vielen Verwendungszusammenhängen an Rechtsprechung fehlt, die die abstrakten Vorgaben ausreichend konkretisieren würde. Es fehlen deshalb konkrete Leitlinien, an denen sich Betroffene, Behörden, die Gesellschaft und Unternehmen orientieren können.

Die Verwaltungsstrukturen der Datenschutzaufsicht verstärken dieses Problem. Es existieren 17 Datenschutzbehörden und nur selten vertreten sie bei kontroversen Themen bis ins Detail die gleichen Positionen. Durch das informelle Gremium »Düsseldorfer Kreis« versuchen die Aufsichtsbehörden zwar, eine gemeinsame Position zu entwickeln. Allerdings sind die resultierenden Hinweise weder verbindlich noch weisen sie den notwendigen Detaillierungsgrad auf; selten sind sie wirklich ausgewogen.

Diese Bedingungen begründen für Unternehmen ein erhebliches ökonomisches Risiko, denn Big-Data-Projekte gehen mit nennenswerten Investitionen einher. Ein Mindestmaß an Rechtssicherheit bezüglich der datenschutzrechtlichen Bewertung ist deshalb unverzichtbar. In vielen entscheidenden Fragen besteht jedoch das Risiko, dass eine Aufsichtsbehörde Maßnahmen als unzulässig erachtet, die das Unternehmen aufgrund einer anderen Auslegung der abstrakten rechtlichen Vorgaben für zulässig hielt.

Grundlegende Zulässigkeitsfragen sind deshalb eindeutig zu klären:

- Wie weit geht der Personenbezug von Daten und wann gelten Daten als anonymisiert<sup>105</sup>?
- Ist die Anonymisierung selbst als Datenverarbeitung zu sehen, die einen gesetzlichen Erlaubnistatbestand erfordert?
- Inwiefern greift die Zweckbindung und welche Befreiungen sind möglich?
- Ist die Verarbeitung von pseudonymisierten Daten unter erleichterten Voraussetzungen zulässig?

## Politischen Diskurs führen – positive Nutzungsregeln entwickeln

Notwendig und wünschenswert ist dabei nicht nur eine rechtliche Präzisierung, sondern auch ein fortwährender gesellschaftlicher und politischer Diskurs, der es ermöglicht, gesellschaftlich akzeptierte Standards zu Big Data zu entwickeln. Dabei besteht allerdings die Gefahr, dass die Debatte sich einseitig auf die Risiken der Technologien konzentriert, ohne die Chancen genügend zu würdigen.<sup>106</sup>

Es ist jedoch genauso unmoralisch und falsch, die Möglichkeit der Heilung von Krankheiten, des Verhinderns von Katastrophen und der Verbesserung des Lebensstandards auszublenden, wie es falsch ist, Risiken zu ignorieren.

Chancen und Risiken müssen stets in einen angemessenen Ausgleich gebracht werden und an diesem Maßstab sollten sich die rechtlichen Vorgaben orientieren. Ganz konkret sollte Big Data nicht – wie bisher – nur unter Schutzaspekten gesetzlich erfasst, sondern positive Nutzungsregeln ins Auge gefasst werden, die der zentralen Bedeutung von Daten in einer Data Driven Economy gerecht werden. Das Recht muss begründete Sorgen und grundrechtliche Positionen berücksichtigen, es darf aber nicht zu einem Institut verkommen, das den gesellschaftlichen Mehrwert technologischer Entwicklungen – und der auch darin zum Ausdruck kommenden Grundrechtsverwirklichungen – ausblendet.

Angesichts der globalen Nutzung von Big-Data-Produkten sind dabei auch die unterschiedlichen kulturellen Sichtweisen zu berücksichtigen. Letztlich erstrebenswert wäre ein globaler Datenschutz, ähnlich einem Weltinformationsethos, wie er bereits auf dem ersten UNESCO-Kongress über ethische und rechtliche Aspekte der digitalen Information im März 1997 in Monaco niedergeschrieben wurde.<sup>107</sup>

---

<sup>105</sup> Mit der Anwendung mathematischer Algorithmen ist es möglich geworden, persönliche Daten zu abstrahieren und in synthetische Daten zu überführen. Diese haben die gleiche Qualität wie die Ursprungsdaten, sind aber auf ihre Herkunft nicht mehr nachzuvollziehen. Diese neuen synthetischen Daten können nun beliebig verwendet, verarbeitet und gespeichert werden. Unternehmen können mit diesen Daten genauso arbeiten wie mit Originaldaten, ohne die Privatsphäre oder Gesetze zu missachten (vgl. (Wegner 2015)).

<sup>106</sup> Diese Tendenz besteht in Teilen der deutschen Medienlandschaft.

<sup>107</sup> Vgl. (INFOethics 1997).

## EU-Datenschutz-Grundverordnung weiterentwickeln

Zum Start der Verhandlungen über die Details der EU-Datenschutz-Grundverordnung forderte Bitkom Nachbesserungen<sup>108</sup>, denn neue Geschäftsmodelle auf der Grundlage von Technologien wie Big Data, Cloud Computing oder Cognitive Computing hängen ganz wesentlich von der konkreten Ausgestaltung der Datenschutz-Grundverordnung ab.

Werden die rechtlichen Grundlagen für die Verarbeitung von persönlichen Daten zu eng gefasst, bleibt zu wenig Spielraum für Innovationen. In diesem Zusammenhang unterstreicht Bitkom, dass eine einseitige Verengung der öffentlichen Diskussion auf Wirtschaft contra Bürgerrechte an den Realitäten vorbei geht.

Bitkom setzt sich für ein ausreichend flexibles Datenschutzrecht ein, das neue Anwendungen ermöglicht, aber die Privatsphäre der Menschen schützt. Besonderes Augenmerk verdienen in der geplanten Verordnung unter anderem die Themen Datensparsamkeit, Zweckbindung, Einwilligung und Profilbildung.<sup>109</sup>

## Rechtliche Anreize schaffen

Ein wesentlicher Schwachpunkt der aktuellen Entwürfe besteht in den hohen Anforderungen an die Anonymisierung bei gleichzeitig fehlenden Anreizen für den Einsatz von Pseudonymisierungstechniken bei der Datenverarbeitung. In der Praxis ermöglicht die Anonymisierung z. B. die Nutzung von Standortdaten von Personen bzw. deren Fahrzeugen für die Verkehrslenkung. Anonymisierte, also irreversible de-personalisierte Daten fallen aus dem Anwendungsbereich der Verordnung heraus, pseudonymisierte, also reversibel de-personalisierte Daten dagegen nicht. Gerade diese sind aber für viele Anwendungen sinnvoll und ermöglichen z. B. eine sichere und datenschutzfreundliche Auswertung der Krankheitsverläufe von Patienten für die medizinische Forschung. Die Analyse von Daten in pseudonymisierter Form führt in vielen Fällen zu einem wesentlich besseren Schutz der Persönlichkeitsrechte des Einzelnen als die Auswertung von Klardaten (mit Einwilligung). Daher müssten in der EU-Datenschutz-Grundverordnung rechtliche Anreize und Erlaubnistatbestände geschaffen werden, damit Unternehmen dieses Element von »Privacy by Design« auch einsetzen (können). Diese könnten darin bestehen, dass die Verarbeitung pseudonymisierter Daten in weiterem Umfang zulässig ist als die Verarbeitung von Klardaten. Setzen Unternehmen oder andere Organisationen die Instrumente der Pseudonymisierung ein, sollte eine Verarbeitung der Daten erleichtert werden.

## Datensparsamkeit und Zweckbindung

Einer Überarbeitung bedürfen auch die Grundprinzipien der Datensparsamkeit und der Zweckbindung bei der Datenerhebung. In einer digital vernetzten Welt spielen Datenvielfalt und

<sup>108</sup> Vgl. (Bitkom, 2015b) sowie (Mühlich/Florian, 2015).

<sup>109</sup> Vgl. (Dehmler, 2015).

Datenreichtum eine entscheidende Rolle. Wirtschaft und Gesellschaft sollten sich daher die Datenvielfalt zunutze machen und dafür sinnvolle Regelungen finden. So ist es das Ziel von Big-Data-Analysen, aus großen Mengen unterschiedlicher Daten neue Erkenntnisse zu gewinnen.

Mit diesem Ziel muss auch der Grundsatz in Einklang gebracht werden, dass personenbezogene Daten nur für einen bestimmten Zweck verarbeitet werden dürfen. Eine Auswertung von Daten sollte zulässig sein, soweit sie für die Betroffenen nicht nachteilig ist bzw. bestehende Risiken für die Rechte Betroffener durch das Ergreifen von gegenläufigen Schutzmaßnahmen (z. B. Pseudonymisierung) adressiert werden. Die bereits im geltenden Recht angelegte grundsätzliche Erlaubnis, Daten für eigene Geschäftszwecke zu verarbeiten oder weiterzuverarbeiten, wenn ein »berechtigtes Interesse« des Unternehmens an der Verarbeitung besteht und dem keine Interessen des Betroffenen entgegenstehen, ist hierfür eine gute Basis. Bei diesem Abwägungsprozess sollte grundsätzlich von einem berechtigten Interesse der Unternehmen bei einer Verarbeitung pseudonymisierter Daten ausgegangen werden.

## Informationspflichten sinnvoll gestalten

Die in der aktuellen Ratsfassung formulierten Informationspflichten bedeuten einen unzumutbaren Aufwand. Die Vorschriften über die Rechte der Betroffenen sollten deshalb mehrstufig angelegt werden. So lassen sich das individuelle Informationsbedürfnis der Betroffenen und die Begrenzung des Aufwandes bei den verantwortlichen Stellen in Einklang bringen. Die proaktiven Informationspflichten sollten sich auf wirklich relevante Informationen beschränken. In einigen Fällen wäre anstelle einer Benachrichtigung eine Hinweispflicht sinnvoll. Das ist zum Beispiel bei der Videoüberwachung der Fall, bei der ein anderes Vorgehen gar nicht möglich ist. Möchte sich der Betroffene daraufhin tiefer informieren, steht ihm ein weitergehendes Auskunftsrecht zu. Im Anschluss an die erteilte Auskunft kann er dann die Rechte auf Berichtigung und Löschung geltend machen.

## Einwilligungen

Einwilligungen sollten dort zum Einsatz kommen, wo kein üblicher Erlaubnistatbestand für die Datenverarbeitung oder ein besonderes Risiko für den Betroffenen besteht. Es darf keine zu starren Formvorgaben geben. Wenn das Instrument Einwilligung zu inflationär gebraucht wird, verliert es seine Warnfunktion, wird lästig und nicht mehr ernst genommen.

Die Informationspflichten sollten weiter entschlackt werden. Schwierig in Bezug auf Big Data wird es insbesondere sein, den Betroffenen so zu informieren, dass er eine informierte Entscheidung treffen kann; bei der Zweitverwertung von Daten wird es oft gar nicht möglich sein, den Betroffenen für eine Einwilligung direkt zu adressieren. In vielen Fällen wie z. B. bei Verwendung von nur pseudonymisiert gespeicherten Daten wird es nicht möglich und auch nicht unbedingt im Sinne des Betroffenen sein, hierfür eine Einwilligung einzuholen. Hier wäre der Erlaubnistatbestand auf Basis einer Interessenabwägung sinnvoll und würde die Erprobung neuer Datenverarbeitungsmethoden unterstützen sowie den administrativen Aufwand begrenzen.

# 10 Autorenverzeichnis

**Prof. Dr. Christoph Bauer** ePrivacy GmbH

**Martin Birkmeier** FIR e. V. an der RWTH Aachen

**Arnd Böken** Graf von Westphalen Rechtsanwälte Partnerschaft

**Dr. Mikio L. Braun** Technische Universität Berlin

**Patrick von Braunmühl** Selbstregulierung Informationswirtschaft e. V.

**Florian Buschbacher** PricewaterhouseCoopers AG Wirtschaftsprüfungsgesellschaft

**Martin Buske** DataXu GmbH

**Susanne Dehmel** Bitkom e. V.

**Guido Falkenberg** Software AG

**Dr. Jörg Friedrichs** Deutsche Telekom AG

**Andreas Hufenstuhl** CSC Deutschland GmbH

**Prof. Dr. habil. Christoph Igel** Deutsches Forschungszentrum für Künstliche Intelligenz

**Dr. Thomas Keil** SAS Institute GmbH

**Ralf Konrad** T-Systems International GmbH

**Holm Landrock** Experton Group AG

**Dr. Mario Lenz** empolis GmbH

**Dr. Ralf Meinberg** Deutsche Telekom AG

**Dr. Flemming Moos**, Osborne Clarke

**Steffen Nienke** FIR e. V. an der RWTH Aachen

**Swen Obermann** Capgemini Deutschland GmbH

**Dr. Jens Schefzig** Osborne Clarke

**Barbara Schmitz** Kabel Deutschland Vertrieb und Service GmbH

**Markus Schröder** Tembit Software GmbH

**Maximilian Thess** Lemma Group UG

**Dr. Carlo Velten** Crisp Research AG

**Dr. Mathias Weber** Bitkom e. V.

**Alexander Widak** Fiducia & GAD IT AG

**Dr. Frank Wisselink** Detecon International GmbH

**Dr. Andreas Zolper** Deutsche Telekom AG

Weiterhin haben mitgewirkt:

**Dr. Dirk Mahnkopf** Cisco Systems GmbH

**Henning Saul** GFT Technologies AG.

# 11 Verzeichnis der Abkürzungen

Abkürzung	Bedeutung
<b>AGB</b>	Allgemeine Geschäftsbedingungen
<b>AutoID</b>	Automatische Identifikation und Datenerfassung oder Automatische Identifizierung
<b>BDSG</b>	Bundesdatenschutzgesetz
<b>Bitkom</b>	Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V.
<b>BMBF</b>	Bundesministerium für Bildung und Forschung
<b>BMWi</b>	Bundesministeriums für Wirtschaft und Energie
<b>CPPS</b>	Cyber Physical Production System
<b>CPS</b>	Cyber Physical System
<b>DRMS</b>	Data Rights Management System
<b>EG</b>	Europäische Gemeinschaft
<b>EKG</b>	Elektrokardiogramm
<b>ETH</b>	Eidgenössische Technische Hochschule
<b>EU</b>	Europäische Union
<b>EuGH</b>	Europäische Gerichtshof
<b>EWR</b>	Europäischer Wirtschaftsraum
<b>GPS</b>	Global Positioning System
<b>HPA</b>	Hamburg Port Authority
<b>IoT</b>	Internet of Things
<b>M2M</b>	Machine-to-Machine Communication
<b>NCT</b>	Nationales Centrum für Tumorerkrankungen
<b>RFID</b>	Radio-frequency Identification
<b>RWTH</b>	Rheinisch-Westfälische Technische Hochschule
<b>SIT</b>	Fraunhofer-Institut für Sichere Informationstechnologie
<b>URL</b>	Uniform Resource Locator
<b>Wifi</b>	Wireless Fidelity
<b>WLAN</b>	Wireless Local Area Network
<b>ZEW</b>	Zentrum für Europäische Wirtschaftsforschung
<b>WLAN</b>	Wireless Local Area Network
<b>ZEW</b>	Zentrum für Europäische Wirtschaftsforschung

## 12 Sachwortverzeichnis

- Adaptivität
  - dynamische 36
  - statistische 36
- Akzeptanz 50, 66
- Algorithmus
  - Ethik 69
  - selbstlernender 37
- Analytics Capability 58
- Anonymisierung 51, 52, 79, 90
- Anonymisierungsprozess 66
- Anreicherung
  - digitale 22
- Apple 98
- Assistenzsystem
  - selbstlernendes 38
- Auskunftsrecht 91
- Auto
  - autonomes 40
- Bankdienstleistung 37
- Behandlungsmuster 51
- Best Practices 72
- Betrugsbekämpfung 38
- Bevölkerung
  - alternde 29
- Bewegungsdaten 24
- Big Brother 71
- Big Data 14
  - internationale Dimension 63
  - positive Nutzungsregel 89
  - Risiken 15
- Big Data Analytics 78
- Big-Data-Anwendung
  - Nutzen 82
- Big-Data-Technologie 15
- Bildungscontrolling 37
- Bildungsnetz
  - Intelligentes 35
- Biodiversität 33
- Bitkom 15, 16, 22, 23, 27, 31, 32, 44, 62, 71, 88, 90, 98, 99, 100, 101, 102
- Bonität 78
- Bonitätsprüfung 78
- Bundesregierung 23
- Bürgerdialog 66
- Business
  - Software-defined 56
- Business Intelligence 37
- Campus-Management-System 36
- Car-to-Car-Communication 40
- Chancengleichheit 17
- Charité 26, 27
- Chemotherapie 27
- Claas 38, 102
- Cloud Computing 23, 29, 56, 90, 97
- Cloud-Lösung 30
- Cloud-Technologie 15
- CO<sub>2</sub>-Emission 24
- Cognitive Computing 32, 57, 62, 88, 90
- Connected Car 22
- Cookies 42
- Customer Journey 65
- Data Driven Economy 18, 88, 89
- Data Economy 57
- Data Mining 74
- Data Rights Management System 76
- Data Scientist 62, 68, 69
- Daten
  - anonymisierte 79, 82
  - besondere 72
  - biometrische 29
  - Eigentum 58
  - Inhaberschaft 74
  - Löschung 75
  - personenbezogene 51, 67
  - pseudonymisierte 67, 82
  - Verwertungsrecht 58
  - Weitergabe 83
- Datenanalysemethode 70
- Datenbestätigung 77
- Datenkultur 58
- Datensammlung
  - automatisierte 14, 62
- Datenschutz 14
  - Datenschutzaufsicht 88
  - Datenschutzbeauftragter 82
  - Datenschutzbehörde 88
  - Datenschutzkonformität 72
  - Datenschutzniveau 65
- Datenschutzrecht 52, 63, 64, 90, 98, 100
  - deutsches 88
- Datensicherheit 69
- Datensparsamkeit 90
- Datenvielfalt 91
- Datenwirtschaft 15, 22, 56, 58
- Deutschland 15, 22, 23, 26, 50, 83, 101
- Diagnostik
  - medizinische 14
- Digital Customer Experience 57
- Digitalisierung 16, 22, 39, 43, 62, 88, 100
- Discovery 30
- Düsseldorfer Kreis 88
- Echtzeit-Bietsystem 42
- Educational Data Mining 36
- Einkaufszentrum 41
- Einwilligung 90
- EKG-Daten 27
- EKG-Signal 29
- Energieversorger 33
- Erbkrankheit 15
- Erdbebenfrüherkennung 34
- Ethik 26, 69
  - Ethik-Kommission 73
- EU-Datenschutz-Grundverordnung 15, 16, 62, 63, 83, 88, 90
- Extremisituation 34
- Extremwetterlage 33
- Fabrik
  - intelligente 44
- Facebook 37, 71
- Finanzdienstleister 37
- Finanzdienstleistung 22
- Finanzkrise 34
- Fitness 30
- Fitnessarmband 29
- Fitnesstracker 29
- Flottenlösung 24
- Flutinformations- und Warnsystem 31
- Forensik 14
- Forschung
  - medizinische 14
- Frühgeborenen-Sterblichkeit 28



- Frühgeborenes [15, 28](#)
- Geldwäsche [38](#)
- Generali [30](#)
- Geschäftsgeheimnis [67, 71](#)
- Geschäftsmodell [15](#)
  - datengetriebenes [56](#)
  - softwaregetriebenes [57](#)
- Gesundheit [51](#)
- Gesundheits-App [30](#)
- Gesundheitsprognose [78](#)
- Gesundheitswesen [15, 22, 26](#)
- Gesundheitszustand [30](#)
- Google [71](#)
- Governance [82](#)
- GPS-Daten [57](#)
- Graphendatenbank [38](#)
- Grundrechtsverwirklichung [89](#)
- Hafenwirtschaft [26](#)
- Hamburg Port Authority [25](#)
- Hamburger Hafen [25](#)
- Handlungsbedarf
  - politischer [17](#)
- Hasso-Plattner-Institut [27](#)
- Herzkrankheit
  - koronare [29](#)
- Hindernis
  - regulatorisches [83](#)
- Hochwasserschutz [14](#)
- Hochwasserschutzzentrale [31](#)
- Indoor Analytics [41](#)
- Industrialisierung
  - datengestützte [39](#)
- Industrie
  - pharmazeutische [28](#)
- Industrie [4.0](#) [15, 22, 42, 43, 44, 62, 88, 100](#)
- Infarktpatient [29](#)
- Information
  - Verfallsdatum [75](#)
- Informations- und Pressefreiheit [77](#)
- Informations- und
  - Wissensdatenbank [36](#)
- Informations- und Wissensplattform [36](#)
- Infrastruktur
  - kritische [32](#)
- Instagram [31](#)
- Intelligenz [44, 70, 102](#)
  - Künstliche [36](#)
- Interessenabwägung [75, 83, 91](#)
- Internet [22](#)
- Internet der Dinge [22, 40, 57](#)
- IT-Sicherheit [23](#)
- Kataster [31](#)
- Katastrophenmanagement [31](#)
- Klimaforschung [34](#)
- Kodex [83](#)
- Korrektiv
  - ethisch-moralisches [72](#)
- Kraftstoffverbrauch [25](#)
- Krankenkasse [30, 51](#)
- Krankenversicherung [78](#)
- Kranker
  - chronisch [15](#)
- Krebs [27](#)
- Krebserkkrankung [15](#)
- Krebsregister [26](#)
- Kreditentscheidung [37](#)
- Kreditprozess [37](#)
- Kriminalitätsvorbeugung [32](#)
- Kundenintegrations-Portal [66](#)
- Künstliche Intelligenz [36](#)
- Landwirtschaft [22, 38](#)
- Learning-Management-System [36](#)
- Leitlinien
  - ethische [15](#)
- Lern- und Forschungsfabrik [35](#)
- Lernermodell [36](#)
  - Lernumgebung personalisierte [35](#)
- Level Playing Field [65](#)
- Logdaten [32](#)
- Logistik [24](#)
- Lokationsdaten [41, 45](#)
- Löschkonzept [76](#)
- Löschrecht [77](#)
- Machine Learning [56](#)
- Makula-Degeneration [28](#)
- Marktregulierung [67](#)
- Medikament
  - personalisiertes [15](#)
- Medizin [66](#)
- Mobilität [22](#)
- Mobilitätsforschung [40](#)
- Mukoviszidose [28](#)
- Mustererkennung [30, 42](#)
- Nahverkehr
  - öffentlicher [25](#)
- Nationales Centrum für Tumorerkrankungen [27](#)
- Naturkatastrophe [34](#)
- Neonatologie [28](#)
- Netflix [57](#)
- Netzwerk
  - Soziales [78](#)
- Nudging [73](#)
- Online-Assessment-Verfahren [36](#)
- Online-Dienst [42](#)
- Online-Marketing [41](#)
- orwellsche Welt [16](#)
- Outdoor Analytics [40](#)
- Politik [15](#)
- Predictive Policing [32](#)
- Privacy by Default [74](#)
- Privacy by Design [74, 90](#)
- Privacy-Preserving Data Mining [52](#)
- Privatsphäre [15, 89, 90](#)
- Product
  - Software-defined [56](#)
- Produktionssystem [44](#)
- Profilbildung [90](#)
- Programmatic Advertising [42](#)
- Prozessmanagement [17, 72](#)
- Pseudonymisierung [90](#)
- Pulsarmband [30](#)
- Quantified Self [30](#)
- Real Time Bidding [42](#)
- Rechtsordnung [64](#)
- Rechtssicherheit [62](#)
- Rechtsvereinheitlichung [64](#)
- Regulierung [17, 83](#)
- Richtlinie 95/46 EG [63](#)

- Routenplanung 25
- Safe Harbor 65
- Sanktionierung 73
- Schweizer Bahn 24
- sd-kama 31, 102
- Seculytics 52
- Selbstbestimmung
  - informationelle 74, 75
- Selbstkontrolle 83
- Selbstregulierung
  - Informationswirtschaft e.V. 92
- Selbstverpflichtung 17, 83
- Senior 15
- Sensorik 57
- Sensor-Technologie 15
- Service-Marktplatz 26
- Siegel 84
- Smart Factory 42
- Smart Grid 23
- Smart Meter 23, 33
- Smart Meter Gateway 33
- Smart Metering 23
- Smart Operation 42
- Smart Port Logistics 25
- Smart Product 42
- Smart Service 42
- smartPORT logistics 26
- Solidargemeinschaft 73
- Solidaritätsprinzip 73
- Sozialleistungen 50
- Start-up 15
- Start-up-Unternehmen 37
- Stromnetz
  - intelligente intelligentes 23
- Swisscom 24
- Telekommunikation 22, 44
- Telekommunikationsinfrastruktur 44
- Tesla 57
- Transparenz 14, 50, 62, 66, 71, 82
  - Dimensionen 67
- Treibstoffverbrauch 24
- Umsatzsteuerkarussell 50
- Umwelt-Monitoring 33
- UNESCO 89
- Unternehmensorganisation 17
- USA 32
- Verbraucherdatenmanagement 57
- Verbrauchercredit 37
- Verbraucherrecht 37
- Verbraucherservice-Cockpit 66
- Verbrechensbekämpfung 66
- Verfallsdatum 75
- Vergessenwerden
  - Recht 16, 63, 75
- Verkehrsplanung 14
- Verkehrssteuerung 39
- Verkehrsüberwachung 14
- Versicherungsbeitrag 30
- Versicherungsbetrug 38
- Versicherungstarif
  - risikoorientierter 73
- Versicherungswesen 73
- Vertrauen 62
- Vertrauensentwicklung 14
- Videoüberwachung 91
- Vitaldaten 29
- Vitalparameter 15, 29, 30
- Weinbau 39
- Weltinformationsethos 89
- Werbeeinblendung 42
- Widerspruchsmöglichkeit 68
- Wildlife Picture Index 33
- Wirtschaft
  - datenbasierte 83
- Wirtschaftswissenschaft 33
- Wissensdienst
  - Intelligenter 36
- Wissensmodell 36
- WLAN 39
- YouTube 31
- Zalando 57
- Zertifikat 84
- Zertifizierung 83
- Zugriffe
  - unberechtigter 82
- Zweckbindung 51, 89, 90

## 13 Quellen

acatech (Herausgeber): Smart Service Welt. Umsetzungsempfehlungen für das Zukunftsprojekt Internetbasierte Dienste für die Wirtschaft. Berlin, März 2014.

Bauer, Christoph: »Big Data und Datenschutz – ein lösbarer Konflikt?«, Datenschutz-Berater, Nr.11, 2013, S. 248–24.

Bauer, Christoph: »Big Data und Datenschutz – geht das?«, Datenschutz-Berater, Nr. 3, 2013, S. 65.

Big Data und Information Governance – Datenflutsteuerung mit Verantwortung.  
<http://www.bigdata-insider.de/recht-sicherheit/articles/465661/> sowie  
<http://www.cloudcomputing-insider.de/plattformen/datenbanken-big-data/articles/471738>

Big Data: Tim Berners-Lee verlangt mehr Datenschutz.  
<http://www.silicon.de/41604472/big-data-tim-berners-lee-verlangt-mehr-datenschutz>

Bitkom Research: Mit Daten Werte schaffen – Report 2015. Studie der Bitkom Research GmbH im Auftrag der KPMG AG Wirtschaftsprüfungsgesellschaft. Juni 2015. Vgl.: <http://www.kpmg.com/DE/de/Documents/mdws-survey-2015-kpmg.pdf> (abgerufen am 08. Juli 2015).

Bitkom, 2012: Big Data im Praxiseinsatz – Szenarien, Beispiele, Effekte. Leitfaden.

Bitkom, 2013: Wie Cloud Computing neue Geschäftsmodelle ermöglicht. Leitfaden.

Bitkom, 2014: Big-Data-Technologien – Wissen für Entscheider. Leitfaden.

Bitkom, 2015: Big Data und Geschäftsmodell-Innovation in der Praxis: 40+ Beispiele. Leitfaden.

Bitkom, 2015a: Kognitive Maschinen – Meilenstein in der Wissensarbeit. Leitfaden.

Bitkom, 2015b: Datenschutz: EU-Datenschutzverordnung muss Innovationen ermöglichen, Presseinformation des Bitkom vom 24. Juni 2015, [http://www.bitkom.org/de/presse/8477\\_82534.aspx](http://www.bitkom.org/de/presse/8477_82534.aspx) (abgerufen am 24. Juni 2015).

Blumtritt, Jörg, Datarella GmbH, BYOD – Bring your own data. Die Chancen und Schwierigkeiten von selbsterhobenen Patientendaten. Vortrag auf dem 1. Meeting des AK Big Data, 29. Januar 2015 (Bitkom-Mitgliederweb).

BMWi, 2015: Vgl. [http://www.digitale-technologien.de/DT/Navigation/DE/Foerderprogramme/Smart\\_Data/Begleitforschung/begleitforschung.html](http://www.digitale-technologien.de/DT/Navigation/DE/Foerderprogramme/Smart_Data/Begleitforschung/begleitforschung.html) (abgerufen am 9. Juli 2015).

Bögershausen, Anne, Hewlett-Packard GmbH: Big Data im Umwelt-Monitoring: Earth Insight – Wildlife Picture Index Analytics System, Vortrag auf dem Meeting der Bitkom-Arbeitskreise Big Data und E-Government, Köln, 27. November 2014.

- Bohsem, Guido: Gesundheits-Apps – Überleben ist wichtiger als Datenschutz. 3. August 2015. Vgl. <http://www.sueddeutsche.de/digital/gesundheit-der-digitale-patient-1.2591921> (abgerufen am 4. August 2015).
- Böken, Arnd, GvW: Leitlinien für den Datenschutz bei Big-Data-Projekten. Vortrag auf dem 1. Big Data Summit (24.06.2013) des Bitkom.
- Bruysten, Tim: Unter der Datenflagge – Big-Data-Schätze im E-Commerce heben. Vortrag auf dem TDWI-Jahreskongress 2015.
- Buschbacher, Florian: Big-Data-Ethik-Kodex: Ziele, Bereitschaft und mögliche Grundsätze. Vortrag auf dem 3. Big Data Summit (25. Februar 2015, [www.bitkom-bigdata.de](http://www.bitkom-bigdata.de)).
- Charta der digitalen Vernetzung, Hamburg, 21. Oktober 2014, Nationaler IT-Gipfel, <http://charta-digitale-vernetzung.de>
- cio.de, 5. August 2015: AOK Nordost – Die Apple Watch auf Rezept. Vgl. [http://www.cio.de/a/die-apple-watch-auf-rezept,3245254?tap=ba569f8fd3d041cb-4c5675eed6a210a5&r=964613837697337&lid=438767&pm\\_In=45](http://www.cio.de/a/die-apple-watch-auf-rezept,3245254?tap=ba569f8fd3d041cb-4c5675eed6a210a5&r=964613837697337&lid=438767&pm_In=45)
- Crisp Research: Datability – Der Markt für datenbasierte Produkte und Services. 2014. Vgl. auch die Presseinformation des Bitkom »Großes Wachstum bei Big Data« vom 05.03.2014: [https://www.bitkom.org/Presse/Presseinformation/Pressemitteilung\\_4293.html](https://www.bitkom.org/Presse/Presseinformation/Pressemitteilung_4293.html) (abgerufen am 7. August 2016).
- Dabrock, Peter: Freiheit, Selbstbestimmung, Selbstkonstruktion – zur Ethik von Big Data. Jahrestagung des Deutschen Ethikrates »Die Vermessung des Menschen – Big Data und Gesundheit«, Berlin, 21. Mai 2015. Vgl. <http://www.ethikrat.org/dateien/pdf/jt-21-05-2015-dabrock.pdf> (abgerufen am 30. August 2015).
- Dammann, Ulrich, In: Simitis, Spiros: Bundesdatenschutzgesetz, 8. Auflage 2014, § 3 Rn. 205 ff.
- Daniel, Kai, RWE: Chancen und Risiken der Verarbeitung großer Datenmengen im Energienetz der Zukunft. Vortrag auf dem 2. Bitkom Big Data Summit, 26. März 2014.
- Data 4 Development – the Orange France Telecom initiative in Ivory Coast. <http://www.d4d.orange.com/home>.
- Dehmel, Susanne: Big Data und europäisches Datenschutzrecht. Bitkom-Positionspapier, 04.02.2015, Vgl.: <https://www.bitkom.org/Bitkom/Publikationen/Big-Data-und-europaisches-Datenschutzrecht.html> (abgerufen am 21. Juli 2015).

Deutsche Bahn 2015: Digitale Revolution bei der DB – Größter Umbruch seit der Bahnreform. Vgl. [http://www.deutschebahn.com/de/presse/themendienste/9103824/medienpaket\\_digitalisierung.html](http://www.deutschebahn.com/de/presse/themendienste/9103824/medienpaket_digitalisierung.html) (abgerufen am 20. August 2015).

Deutsche Telekom: Leitsätze der Deutschen Telekom zu Big Data. Vgl. <https://www.telekom.com/static/-/205806/2/big-data-leitlinien-si> (abgerufen am 26. August 2015).

DIN 66398 – Leitlinie zur Entwicklung eines Löschkonzepts mit Ableitung von Löschfristen für personenbezogene Daten. Vgl.: <http://www.nia.din.de/cmd?level=tpl-art-detailansicht&committeeid=54738935&artid=227394449&languageid=de&bcrumplevel=2>

Doing good with analytics: the pledge. <https://duncan3ross.wordpress.com/2012/11/03/doing-good-with-analytics-the-pledge>

Doing good with data. <http://duncan3ross.wordpress.com/2012/10/25/doing-good-with-data-the-case-for-the-ethical-data-scientist>

Dornacher, Christian, Hitachi Data Systems GmbH: Shinkansen – Oder warum Züge pünktlich kommen (können)! Vortrag auf dem 3. Bitkom Big Data Summit, 25. Februar 2015.

Dörner, Stephan: Krankenkassen bezuschussen die Apple Watch. In: Die Welt, 6. August 2015, S. 10.

Dörner, Stephan; Seibel, Karsten: Prüft Facebook die Bonität? Der Freundeskreis könnte über die Kreditwürdigkeit entscheiden. In: Die Welt, 7. August 2015, S. 15.

Drilling, Thomas; Litzel, Nico: Ein Netzwerk des Guten schaffen – Goethe Universität fördert Big-Data-Nutzung zum Wohl der Allgemeinheit. 23. Juli 2015. Vgl. <http://www.bigdata-insider.de/analytics/articles/498342/?cmp=nl-274> (abgerufen am 4. August 2015).

Eckert, Klaus-Peter; Henckel, Lutz; Hoepner, Petra: Big Data – ungehobene Schätze oder digitaler Albtraum. Fraunhofer FOKUS, März 2014.

Engeser, Manfred: Die Macht der Daten schlägt in allen Branchen durch. Vgl.: URL: <http://www.cio.de/2958578> (abgerufen am 23. Juli 2014).

Ernst & Young GmbH: Kosten-Nutzen-Analyse für einen flächendeckenden Einsatz intelligenter Zähler. Endbericht zur Studie im Auftrag des Bundesministeriums für Wirtschaft und Technologie, Juli 2013. Vgl. <http://www.bmwi.de/DE/Mediathek/publikationen,did=586064.html>

EuGH: Urteil des Gerichtshofs (Große Kammer) vom 13. Mai 2014. Vgl.: <http://curia.europa.eu/juris/document/document.jsf?text=&docid=152065&pageIndex=0&doclang=DE&mode=req&dir=&occ=first&part=1> (abgerufen am 31. Juli 2015).

- Europäische Kommission, 2. Juli 2014: Für eine florierende datengesteuerte Wirtschaft. Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen. Brüssel, COM(2014) 442 final.
- European Commission, ohne Jahresangabe: [http://ec.europa.eu/information\\_society/news-room/cf/dae/document.cfm?doc\\_id=1628](http://ec.europa.eu/information_society/news-room/cf/dae/document.cfm?doc_id=1628) (abgerufen am 28. Juli 2015).
- Graubner-Müller, Alexander, Kreditech Holding SSL GmbH: Einsatz von Big Data in der »Consumer Online Finance«-Industrie. Vortrag auf dem 3. Bitkom Big Data Summit, 25. Februar 2015.
- Grzanna, Marcel: Der Orwell'sche Bürger. 02.07.2015, vgl. <http://www.heise.de/tr/artikel/Der-Orwell-sche-Buerger-2663947.html> (abgerufen am 20. August 2015).
- Handelsblatt Research Institute (im Auftrag der Deutschen Telekom), März 2015: Datenschutz und Big Data – Ein Leitfaden für Kunden. (<http://www.telekom.com/medien/konzern/270672>).
- Handelsblatt Research Institute (im Auftrag der Deutschen Telekom), März 2015: Datenschutz und Big Data – Ein Leitfaden für Unternehmen. (<http://www.telekom.com/medien/konzern/270672>).
- Heise online, 24. Januar 2011, »Digitaler Radiergummi« ist gestartet. Vgl. <http://www.heise.de/newsticker/meldung/Digitaler-Radiergummi-ist-gestartet-1175979.html> (abgerufen am 31. Juli 2015).
- Hofstetter, Yvonne: Big Data und die Macht des Marktes. Aus Politik und Zeitgeschichte, 65. Jg, 11-12/2015, 09. März 2015, S. 33–38.
- INFOethics 1997: First International Congress on Ethical, Legal, and Societal Aspects of Digital Information. Vgl. [http://portal.unesco.org/ci/en/ev.php-URL\\_ID=5379&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/ci/en/ev.php-URL_ID=5379&URL_DO=DO_TOPIC&URL_SECTION=201.html) (abgerufen am 25. August 2015).
- Informationsethik und digitale Gesellschaft. <http://www.forum-privatheit.de/forum-privatheit-de/aktuelles/veranstaltungen/veranstaltungshinweise/vergangene-veranstaltungen/ringvorlesung-informationsethik-und-digitale-gesellschaft.php>
- Jakobs, Joachim; Litzel, Nico: Übernimmt künstliche Intelligenz die Steuerung? efahren von Big Data, der Digitalisierung und Industrie 4.0, Teil 2. 20. Februar 2015, Vgl. <http://www.bigdata-insider.de/recht-sicherheit/articles/475100/?cmp=nl-274>
- Jakobs, Joachim; Litzel, Nico: Viele Daten, viele Risiken? Gefahren von Big Data, der Digitalisierung und Industrie 4.0, Teil 1. 28. Januar 2015. Vgl. <http://www.bigdata-insider.de/recht-sicherheit/articles/472572>

Juran, Sabrina, United Nations Population Fund (UNFPA); Gerland, Patrick, United Nations Department of Economic and Social Affairs (UNDESA), Population Division: Mobilizing a Data Revolution for Sustainable Development.  
<http://paa2015.princeton.edu/uploads/153811> (abgerufen am 15. Juni 2015).

Kemeröz, Tamer, media transfer AG; Kremer, Werner, T-Systems International GmbH: IT Sicherheit in kritischen Infrastrukturen am Praxisbeispiel des Deutschen Smart Meter Roll-Outs. Vortrag auf dem 3. Bitkom Big Data Summit, 25.02.2015.

Kempf, Dieter: »Daten sind ein viel zu kostbares Gut, um sie ungenutzt liegen zu lassen«, Keynote auf dem Safer Internet Day (10. Februar 2015). Vgl. <http://www.digitalewelt.org/content/daten-sind-ein-viel-zu-kostbares-gut-um-sie-ungenutzt-liegen-zu-lassen>.

Kempf, Dieter: Sind Daten die Währung von morgen? Rede im ZEW Mannheim, 2. März 2015.

Klemm, Balázs, Almax Capital GmbH & Co. KG; Eckl, Horst, Clueda AG, München: Visualisierung des Risikonetzwerkes im Finanzdienstleistungsbereich. Vortrag auf dem 3. Bitkom Big Data Summit, 25. Februar 2015.

Knowledge@Detecon (2015): Future Telco Reloaded – Strategien für eine erfolgreiche Positionierung im Wettbewerb, DETECON Consulting, Köln 2015.

Koelwel, Dunja: Feiner als das Bauchgefühl: Polizei will künftige Tatorte mittels Software voraussagen. Vgl. <http://xethix.com/feiner-als-das-bauchgefuehl-polizei-will-kuenftige-tatorte-mittels-software-voraussagen>

Kreibich, Rolf: Von Big zu Smart – zu Sustainable? Aus Politik und Zeitgeschichte, 65. Jg, 11–12/2015, 9. März 2015, S. 20–26.

Kühling, Jürgen: Big Data: Aufgeben oder Aufgabe? Vgl. <http://www.faz.net/aktuell/politik/die-gegenwart/datenschutz-big-data-aufgeben-oder-aufgabe-13335705.html> (abgerufen am 28. Juli 2015).

Landrock, Holm; Schonschek, Oliver; Gadatsch, Andreas: Big Data Vendor Benchmark 2015 – Big-Data-Lösungsanbieter im Vergleich. Eine Untersuchung der Experton Group AG. München, September 2014.

Langkafel, Peter: Dr. Algorithmus? Big Data in der Medizin. Aus Politik und Zeitgeschichte, 65. Jg, 11–12/2015, 9. März 2015, S. 27–32.

Lixenfeld, Christoph (21. Februar 2014). Big Data in Echtzeit – Lokführer bekommt Zug-Tempo aufs iPad, <http://www.cio.de/a/lokfuehrer-bekommt-zug-tempo-aufs-ipad,2947072>

- Mayer-Schönberger, Viktor: Delete: Die Tugend des Vergessens in digitalen Zeiten. University Press, Berlin 2010.
- Mayer-Schönberger, Viktor: Zur Beschleunigung menschlicher Erkenntnis. Aus Politik und Zeitgeschichte, 65. Jg, 11–12/2015, 9. März 2015, S. 14–19.
- Menschenwürde untergraben – Hatte Orwell mit Big Brother doch recht?  
<http://www.cio.de/a/hatte-orwell-mit-big-brother-doch-recht,3100603>
- Moos, Flemming, Osborne Clarke: Regeln für die Nutzung identitätsrelevanter Daten bei Big Data: welche Anforderungen stellen Datenschutzrecht und Datenschutzaufsichtsbehörden? Vortrag auf dem 1. Meeting des AK Big Data, Januar 2015.
- Morozov, Evgeny: »Ich habe doch nichts zu verbergen«. Aus Politik und Zeitgeschichte, 65. Jg, 11–12/2015, 9. März 2015, S. 3–7.
- Mühlich, Regina; Stocker Florian: Datenschutz in der Industrie 4.0. 24. Juni 2015, <http://www.bigdata-insider.de/recht-sicherheit/articles/492000/> (abgerufen am 24. Juni 2015).
- Niemann, Fabian: 23 Fragen zu Big Data und Recht, Checkliste, Vgl.:  
<https://www.artegic.de/big-data-und-recht> (abgerufen am 15. Juni 2015).
- Nold, Andreas: Am Puls der Zeit – Analyse Sozialer Netzwerke. In: Vitako aktuell, 1/2013, S. 16–17.
- Petersberger Erklärung: Digitalisierung der Gesellschaft – Selber gestalten statt sich überrollen zu lassen! <http://www.comma-soft.com/veranstaltungen/petersberger-gespraech/petersberger-erklaerung>
- Petri, Thomas: Datenflut und Datenschutz – Rechtsfragen. Jahrestagung des Deutschen Ethikrates »Die Vermessung des Menschen – Big Data und Gesundheit«, Berlin, 21. Mai 2015. Vgl. <http://www.ethikrat.org/dateien/pdf/jt-21-05-2015-petri.pdf> (abgerufen am 30. August 2015).
- PIK (Potsdamer Institut für Klimafolgenforschung, 2014): <https://www.pik-potsdam.de/news/press-releases/archive/2014/new-forecasting-method-predicting-extreme-floods-in-the-andes-mountains> (abgerufen am 2. Juli 2015).
- Pluta, Werner (02.04.2008): Interview: »Daten brauchen ein Verfallsdatum«. Golem.de im Gespräch mit Harvard-Professor Viktor Mayer-Schönberger. Vgl. <http://www.golem.de/0804/58721.html> (abgerufen am 31. Juli 2015).
- Porzig, Alf; Vospeter, Philip: Group Wide Machine Data Analytics @ CLAAS. Vortrag auf dem 3. Bitkom Big Data Summit 2015, Vgl. <http://www.bitkom-bigdata.de/programm/2015-02-25/group-wide-machine-data-analytics-claas> (abgerufen am 29. Juli 2015).



Projektgruppe Smart Data der Arbeitsgruppe 2 »Vernetzte Anwendungen und Plattformen für die digitale Gesellschaft«: Smart Data – Potenziale und Herausforderungen. Strategiepapier für den Nationalen IT Gipfel 2014.

Reißmann, Ole: Recht auf Vergessen: Google entfernt SPIEGEL-Artikel aus Suchergebnissen. 04.07.2014. Vgl. <http://www.spiegel.de/netzwelt/netzpolitik/recht-auf-vergessen-google-entfernt-spiegel-artikel-aus-treffern-a-979255.html#ref=rss> (abgerufen am 31. Juli 2015).

Roland Berger Strategy Consultants/BDI: Die digitale Transformation der Industrie. Eine europäische Studie von Roland Berger Strategy Consultants im Auftrag des BDI, Berlin 2015.

Rülke, Kai: Geschwindigkeit ist keine Hexerei – viele Einzelmaßnahmen ergeben eine leistungsfähige Big-Data-Anwendung. In: Vitako aktuell, 1/2013, S. 12–13.

Schallaböck, Jan, iRights.Lab (o.J.): Big Data – Themenpapier zum Projekt »Braucht Deutschland einen Digitalen Kodex«.

Schlegel, Victor, Swisscom (Schweiz) AG: Mobility Insights. Vortrag auf dem 3. Bitkom Big Data Summit, 25. Februar 2015.

Schreier, Jürgen; Litzel, Nico: Großprojekt zu Big Data in der Produktion gestartet. 7. Oktober 2014, vgl. <https://www.bigdata-insider.de/industrie40/articles/461865> (abgerufen am 22. Juli 2015).

Simon, Anja, Deutsche Lufthansa AG: Big Analytics: Fliegen nach Plan – Streckennetzplanung und Marktprognose bei der Lufthansa. Keynote auf dem 2. Bitkom Big Data Summit, 26. März 2014.

Smolan, Rick; Erwitte Jennifer (Herausgeber): The Human Face of Big Data. AAO Against All Odds Productions, 2013.

Steinebach Martin, Fraunhofer SIT: Ergebnisse des Bürgerdialogs Big Data. Vortrag im Bitkom-Arbeitskreis Big Data, 29. Januar 2015.

Steinebach, Martin; Winter, Christian; Halvani, Oren; Schäfer, Marcel; Yannikos, York: Chancen durch Big Data und die Frage des Privatsphärenschutzes. Begleitpapier Bürgerdialog, Fraunhofer-Institut für Sichere Informationstechnologie. SIT Technical Reports SIT-TR-2015-06, März 2015. [https://www.sit.fraunhofer.de/fileadmin/dokumente/studien\\_und\\_technical\\_reports/Big-Data-Studie2015\\_FraunhoferSIT.pdf](https://www.sit.fraunhofer.de/fileadmin/dokumente/studien_und_technical_reports/Big-Data-Studie2015_FraunhoferSIT.pdf) (abgerufen am 8. Juli 2015).

Stolberg, Martin: Analytics für Compliance – In-Memory erkennt Muster von Betrügern. 24. Juni 2015, Vgl. [http://www.cio.de/a/in-memory-erkennt-muster-von-betruergern,3109322?tap=ba569f8fd3d041cb4c5675eed6a210a5&r=964612932997338&lid=429298&pm\\_In=56](http://www.cio.de/a/in-memory-erkennt-muster-von-betruergern,3109322?tap=ba569f8fd3d041cb4c5675eed6a210a5&r=964612932997338&lid=429298&pm_In=56) (Abruf am 24.06.2015)(abgerufen am 26. August 2015).Ulbricht, Carsten: Schwieriges Terrain – Big Data ist eine Herausforderung für den Datenschutz. In: Vitako aktuell, 1/2013, S. 14–15.

- TDWI E-Book: The Internet of Things: The Practical and Ethical Dimensions of Big Data. August 2015.
- Ulmer, Claus, Deutsche Telekom AG: Big Data – Große Datenmengen – Große Sorgen?, Vortrag auf dem 1. Big Data Summit (24. Juni 2013) des Bitkom.
- Ulmer, Claus: Big Data – neue Geschäftsmodelle, neue Verantwortlichkeiten? RDV – Recht der Datenverarbeitung (2013), S. 225 (229).
- Umbach, Christian, Lufthansa Innovation Hub Berlin; Kellner, Stefan, Door2Door GmbH: Urban Transit Technology Solutions durch Big Data: Das Projekt Time2Gate. Vortrag auf dem 3. Bitkom Big Data Summit, 25. Februar 2015.
- Vospeter, Philip, Claas KGaA mbH; Porzig, Alf, Mieschke Hofmann und Partner Gesellschaft für Management- und IT-Beratung mbH Group Wide Machine Data Analytics @ CLAAS. Vortrag auf dem 3. Bitkom Big Data Summit, 25. Februar 2015.
- Wagner, Ben: The Ethics of Algorithms: from radical content to self-driving cars. Final Draft Background Paper. Global Conference on CyberSpace 2015, 16–17 April 2015, The Hague, The Netherlands. Vgl. [https://www.gccs2015.com/sites/default/files/documents/Ethics\\_Algorithms-final%20doc.pdf](https://www.gccs2015.com/sites/default/files/documents/Ethics_Algorithms-final%20doc.pdf) sowie <https://www.gccs2015.com/programme> und <https://cihr.eu/the-ethics-of-algorithms>
- We harness the power of data science in the service of humanity. <http://www.datakind.org>
- Weber, Norbert, Motionlogic GmbH: Big Data im öffentlichen Nahverkehr – mit Aktivitätsdaten aus dem Mobilfunknetz Streckennetze optimieren. Vortrag auf dem 3. Bitkom Big Data Summit, 25. Februar 2015.
- Wegner, Susan, Deutsche Telekom AG: Future Analytics – Fabrizieren von synthetischen Daten. Vortrag auf dem 3. Bitkom Big Data Summit, 25. Februar 2015.
- Weichert, Thilo: Big Data und Datenschutz, <https://www.datenschutzzentrum.de/bigdata/20130318-bigdata-und-datenschutz.pdf> (Zugriff am 15. Juni 2015).
- Welchering, Peter: Digitaler Radiergummi – Was im Netz steht, bleibt. 14. Januar 2011. Vgl. <http://www.faz.net/aktuell/technik-motor/computer-internet/digitaler-radiergummi-was-im-netz-steht-bleibt-1580048.html> (abgerufen am 31. Juli 2015).
- Westermann, Sascha, Hamburg Port Authority AÖR; Müller, Hartmut, T-Systems International GmbH: Smart Port Logistics im Hamburger Hafen, Vortrag auf dem 2. Bitkom Big Data Summit, 26. März 2014.

Willkomm, Marlene, Stadtentwässerungsbetriebe Köln: Smart Data Katastrophenmanagement (sd-kama) – am Beispiel der Hochwasserschutzzentrale der Stadt Köln, Vortrag auf dem Meeting der Bitkom-Arbeitskreise Big Data und E-Government, Köln, 27. November 2014.

Winter, Christian, Fraunhofer-SIT; Mollat, Ingolf, Ernst & Young GmbH; Steinebach Martin, Fraunhofer-SIT: Betrugserkennung in Buchungsdaten – Forensik in Big Data mittels modellbasierter Ziffernanalyse. Vortrag auf dem 2. Bitkom Big Data Summit, 26. März 2014.

Xu, Feiyu, Deutsches Forschungszentrum für Künstliche Intelligenz (und weitere 12 Autoren): Smart Data For Mobility (SD4M) – eine Big-Data-Analytik-Plattform für multimodale Smart Mobility Services. Vortrag auf dem 3. Bitkom Big Data Summit, 25. Februar 2015.

Bitkom vertritt mehr als 2.300 Unternehmen der digitalen Wirtschaft, davon gut 1.500 Direktmitglieder. Sie erzielen mit 700.000 Beschäftigten jährlich Inlandsumsätze von 140 Milliarden Euro und stehen für Exporte von weiteren 50 Milliarden Euro. Zu den Mitgliedern zählen 1.000 Mittelständler, 300 Start-ups und nahezu alle Global Player. Sie bieten Software, IT-Services, Telekommunikations- oder Internetdienste an, stellen Hardware oder Consumer Electronics her, sind im Bereich der digitalen Medien oder der Netzwirtschaft tätig oder in anderer Weise Teil der digitalen Wirtschaft. 78 Prozent der Unternehmen haben ihren Hauptsitz in Deutschland, 9 Prozent kommen aus Europa, 9 Prozent aus den USA und 4 Prozent aus anderen Regionen. Bitkom setzt sich insbesondere für eine innovative Wirtschaftspolitik, eine Modernisierung des Bildungssystems und eine zukunftsorientierte Netzpolitik ein.

**Bundesverband Informationswirtschaft,  
Telekommunikation und neue Medien e.V.**

Albrechtstraße 10  
10117 Berlin  
**T** 030 27576-0  
**F** 030 27576-400  
bitkom@bitkom.org  
[www.bitkom.org](http://www.bitkom.org)

**bitkom**