

Leitlinien für den Big-Data-Einsatz im Überblick Chancen und Verantwortung

Positionspapier der Projektgruppe Smart Data
in der Fokusgruppe Intelligente Vernetzung
Plattform „Innovative Digitalisierung der Wirtschaft“ im Nationalen IT-Gipfel

Inhalt

1. Einleitung	4
2. Nutzen von Big-Data-Anwendungen für Gesellschaft und Privatpersonen	5
2.1 Energiewende	5
2.2 Verkehrssysteme	5
2.3 Medizinische Forschung und Diagnostik	6
2.4 Öffentlicher Bereich	8
2.5 Vorhersage von Krisensituationen	8
2.6 Bildung und Qualifizierung	8
2.7 Finanzdienstleistungen	9
2.8 Landwirtschaft	9
2.9 Connected Car – aus Sicht privater Nutzer	9
2.10 Digitale Einkaufswelten	10
2.11 Online-Marketing und Nutzen von Big Data für Verbraucher	10
2.12 Industrie 4.0	11
2.13 Telekommunikation	11
2.14 Verbesserung des Datenschutzes in Rechenzentren	11
3. Bisher ungenutzte Chancen beim Einsatz von Big Data	12
4. Bedeutung der Datenwirtschaft in der zukünftigen Wirtschaft	14
5. Fragestellungen für Big-Data-Leitlinien	16
5.1 Big Data im internationalen Wettbewerb	16
5.2 Vertrauensentwicklung bei Verbrauchern - Best Practices	17
5.3 Transparenz bei Big-Data-Analysen	17
5.4 Wie Analyseergebnisse zustande kommen	18
5.5 Bedeutung ethischer Grundsätze bei Big-Data-Lösungen	18
5.6 Stimulierung gesellschaftlich wünschenswerter Verhaltensweisen	20
5.7 Datensammlung im Automatismus	20
5.8 Recht auf Vergessenwerden?	21
5.9 Umgang mit Wahrscheinlichkeiten und Verhaltensprognosen	22
5.10 Erkenntnisse aus der Kommunikation in sozialen Netzwerken	23
6. Leitlinien für den Big-Data-Einsatz	24
7. Big Data und Datenschutz: Politischer Handlungsbedarf	27
8. Autorenverzeichnis	30
9. Mitglieder Projektgruppe Smart Data	31

1. Einleitung

Ziele

Diese Publikation¹ verfolgt das Ziel, rechtspolitische Positionen und Leitlinien für den ethisch fundierten, verantwortungsbewussten Einsatz von Big Data zu formulieren. Sie soll auch zu mehr Transparenz im Bereich Big-Data-Anwendungen beitragen, denn Information und Transparenz schaffen die Basis für Vertrauen.

Zielgruppen

Das Positionspapier adressiert Entscheidungsträger in der Politik und in den Datenschutzbehörden, private Verbraucher sowie die Öffentlichkeit. Es wendet sich ebenfalls an die Nutzer von Big-Data-Technologien und -Lösungen in der Wirtschaft, deren Anbieter sowie Ethikbeauftragte.

Schwerpunkte

Das Positionspapier konzentriert sich auf zwei Schwerpunkte:

- Mit Beispielanwendungen aus vielen Bereichen des gesellschaftlichen Lebens wird der vielfältige Nutzen für die Verbraucher und die Gesellschaft insgesamt gezeigt, den Big-Data-Anwendungen mit sich bringen.
- Den zweiten Schwerpunkt bilden Positionen zu Fragestellungen, die im Zusammenhang mit Big-Data-Anwendungen im Fokus des öffentlichen Interesses stehen. Aus diesen Erörterungen werden dann Leitlinien für den Big-Data-Einsatz und Vorschläge an die Politik abgeleitet.

¹ Basis ist das Positionspapier „Leitlinien für den Big-Data-Einsatz – Chancen und Verantwortung“ des Bitkom vom September 2015 (<https://www.bitkom.org/Bitkom/Publikationen/Leitlinien-für-den-Big-Data-Einsatz.html>). Es enthält ein umfangreiches Quellenverzeichnis; daher wird an dieser Stelle auf Quellenangaben verzichtet.

2. Nutzen von Big-Data-Anwendungen für Gesellschaft und Privatpersonen

Das Kapitel 2 zeigt an Beispielen aus 14 ausgewählten Bereichen die Bedeutung von Big-Data-Technologien für die Gesellschaft und die Verbraucher auf.

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. 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.

2.1 Energiewende

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 flächendeckenden Einsatz von Smart Metering in Deutschland vorgestellt. 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.

2.2 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

Die Schweizer Bahnen regeln die Geschwindigkeit ihrer Züge mit neuen Technologien. Die Verminderung von Verkehrshalten lässt auch den Energieaufwand sinken – mit Millioneneinsparungen in jedem Jahr und Beiträgen zum Umweltschutz. 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. Auch die Deutsche Bahn setzt eine umfassende Digitalisierungsstrategie um.

Umweltfreundliches Truckmanagement

Ein global tätiges deutsches Logistikunternehmen hat eine Flottenlösung eingeführt, die zu signifikanten Reduktionen des Treibstoffverbrauchs und der CO₂-Emissionen sowie gleichzeitig zu einem verbesserten Überblick für Disponenten führt. 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 dadurch die Berechnung der Fahreffizienz von Trucks.

Bessere Angebote für den öffentlichen Nahverkehr

Vergleichbare Anwendungen entwickeln auch die Telekommunikations- und andere Unternehmen: 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.

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

Um die Verkehrsströme künftig miteinander zu vernetzen und effizienter zu gestalten, führte die Hamburg Port Authority (HPA) mit IT-Dienstleistern das Cloud-basierende Informations- und Kommunikationssystem smartPORT logistics

ein, das den Andrang der täglich über 140.000 Lkw-Fahrten im Hafen entzerrt und die Güterströme optimiert.

Die Menge an Verkehrs- und Transportdaten ist riesig. Dabei gibt es viele verschiedene Beteiligte wie Hafenverwaltung, Speditionen, Transporteure, Parkflächen- und Terminal-Betreiber. smartPORT logistics ermöglicht es Häfen sowie Transport- und Logistikunternehmen, Transportaufträge in Echtzeit zu überwachen.

2.3 Medizinische Forschung und Diagnostik

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

Es entstehen derzeit Falldatenbanken – z. B. ein Krebsregister an der Charité – für einige häufige sowie für schwere Erkrankungen, um bei neuen Patienten mit ähnlichen Symptomen möglichst wirksame Therapien zu finden.

Die im Abschnitt 2.3 skizzierten Beispielanwendungen 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.

Diagnose und Behandlung von Tumoren

Anfang Juli 2015 stellten das Potsdamer Hasso-Plattner-Institut und die Berliner Charité 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.

Auch am Nationalen Centrum für Tumorerkrankungen (NCT) in Heidelberg arbeiten Ärzte und Wissenschaftler an neuen Therapieformen.

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, auf ihren Fall zugeschnittene Behandlung erhalten. Dafür muss das NCT mit Bezug auf das Patientenprofil eine große Menge an Daten in Echtzeit auswerten. Big-Data-Technologien integrieren all diese Daten mit sehr großen Datenvolumen aus unterschiedlichen Quellen. Die beschleunigte Tumordiagnostik kann die Lebenserwartung der Tumorkranken 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 Gesundheitsprozesses 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, innerhalb eines Jahres einem erneuten Infarkt zu erliegen. So konnten neue Techniken für die Risikoanalyse entwickelt werden.

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 Big-Data-Methoden erlauben es, personalisierte Arzneimittel herzustellen und betroffenen Menschen ein deutlich längeres und besseres Leben zu ermöglichen.

Frühgeborenen-Überwachung

Kanadische Neonatologen haben festgestellt, dass es Fälle von Infektionen bei Frühgeborenen gibt, die durch medizinische Frühindikation 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.

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 bei der japanischen und britischen Bahn zur Optimierung der Verfügbarkeit, Pünktlichkeit und Auslastung der Züge eingesetzt. Auch die Deutsche Bahn setzt eine umfassende Digitalisierungsstrategie um.

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 (Fitnessstracker) gesteigert, die leicht mit dem Smartphone zu koppeln sind. Das Smartphone wird heute durch entsprechende Apps für Sportübungen, Lauftraining oder täglichen Gesundheitscheck zum persönlichen Kontrollzentrum und zur Steuerzentrale für den Einzelnen. 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 Nutzungsprofile und zehntausender Datenpunkte für einzelne Nutzer ermöglicht die maschinelle Mustererkennung als Big-Data-Anwendung. So können Signale für Erkrankungen und Früherkennung von gesundheitlichen Gefährdungssituationen genauer als je zuvor wahrgenommen werden.

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. Einige Anbieter erwägen, Versicherungsbeiträge zu reduzieren.

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.

2.4 Öffentlicher Bereich

Im öffentlichen Bereich können Big-Data-Technologien bei der Erfüllung vieler öffentlicher Aufgaben helfen, wie etwa der Daseinsvorsorge, Umwelt-Monitoring, Stadtentwicklung und öffentliche Sicherheit (Sicherheit kritischer Infrastrukturen). Großes Potenzial bieten Wirtschaftsförderung und Smart-City-Ansätze. Möglich sind auch Anwendungen zur Abschätzung von Gesetzesfolgen oder zur Meinungsanalyse. Vor dem breiteren Einsatz von Big Data im öffentlichen Sektor braucht es klare rechtliche und ethische Grenzen.

Katastrophenmanagement

Die Hochwasserschutzzentrale der Stadt Köln setzt mit Partnern aus Wissenschaft und Wirtschaft das Projekt Smart-Data-Katastrophenmanagement (sd-kama) um. Ziel des Projektes sd-kama ist es, eine Informationsplattform aufzubauen, die ein effizientes Katastrophenmanagement ermöglicht. 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.

Cybersecurity

Einen wichtigen Beitrag zur Sicherheit öffentlicher IT-Infrastrukturen leistet die Überwachung der im IT-Betrieb generierten und mit Big-Data-Ansätzen ausgewerteten Maschinendaten. 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.

Die Polizei prognostiziert mit Big-Data-Analysen Tatorte und Delikte. Auf Basis gesammelter historischer Daten entwickelt sie dann Gegenstrategien.

2.5 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 in den Wirtschaftswissenschaften bei der Vorhersage von Finanzkrisen. Fortschritte in der Wissenschaft ermöglichen es Politik, Wirtschaft und Gesellschaft, früher auf zu erwartende Extremsituationen zu reagieren und dadurch gewaltige Schäden abzuwenden.

2.6 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 marktge-
rechte, nachfrageinduzierte Realisierung von Unterrichts-,
Studien- und Weiterbildungsangeboten für das lebenslan-
ge Lernen ermöglicht.

2.7 Finanzdienstleistungen

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

Einsatz von Big Data bei der Kreditbeschaffung

Im traditionellen Bankgeschäft werden seit vielen Jahren
analytische Verfahren eingesetzt. Diese erleichtern und ob-
jektivieren Kreditentscheidungen etwa beim Hauskauf oder
im Bereich der Verbraucherkredite. Big-Data-Analysen er-
lauben nun wesentlich genauere und schnellere Entschei-
dungen. 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, son-
dern geben auf einer viel breiteren Basis ein genaueres
Bild über das voraussichtlich Zahlungsverhalten und damit
des Risikos, das die Bank eingeht.

Aufdeckung von Versicherungsbetrug

Versicherungsbetrug und Geldwäsche locken raffiniert
vorgehende kriminelle Kreise an, die sich häufig sehr wirk-
sam der Aufdeckung ihrer Betrügereien entziehen. Daraus
erwächst die Herausforderung, durch Netzwerk-Analyse
Zusammenhänge bei Finanztransaktionen in Echtzeit zu er-
kennen. Hier kommen spezielle Datenbanken 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 Missbrauchsschä-
den und können so ihre Einnahmen verbessern bzw. die

Preise für ihre Dienstleistungen senken und damit ihren
Kunden attraktivere Angebote unterbreiten.

2.8 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.

Als fortgeschrittener Nutzer von Big-Data-Technologi-
en statten Landmaschinenhersteller ihre Produkte mit
Sensoren und Messsystemen aus, die einen Datenstrom
erzeugen. Die Daten aus dem Entwicklungs- und Produk-
tionsprozess werden durch Telematics- und Betriebsdaten
sowie Daten interner Systeme für die Ersatzteilbestellung
oder Schadensmeldungsbearbeitung ergänzt. Aus der
Erweiterung des zentralen Datenpools durch Zusatzinfor-
mationen wie z. B. Wetterhistorien oder Bodenbeschaf-
fenheitsdaten entsteht ein Data Lake² mit umfassenden
Big-Data-Einsatzszenarien. Big-Data-Technologien werden
auch zur Optimierung von Wasser- und Düngerzufuhr für
landwirtschaftliche Kulturen eingesetzt.

2.9 Connected Car – aus Sicht privater Nutzer

Sensor-, Cloud- und Big-Data-Technologien bilden die
Grundlage für autonome Fahrzeuge und lassen tiefgrei-
fende Veränderungen in der Ausgestaltung von Mobilitäts-
lösungen erwarten. Zurzeit testen führende Automobil-
hersteller Modelle, die in wenigen Jahren auf den Markt
kommen werden. Die futuristischen Fahrzeuge lenken,
fahren und bremsen allein, gesteuert von Kameras, Sen-
soren 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 Verkehrsflus-

2 Vgl. https://en.wikipedia.org/wiki/Data_lake

ses. Das revolutioniert die Verkehrssteuerung. Das Auto wandelt sich stärker zum Arbeits- oder Lebensraum. Neben dem Austausch von Informationen zwischen zentralen Stellen und den Fahrzeugen 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.

2.10 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 – im Interesse des Einzelnen und der Gesellschaft gleichermaßen.

Outdoor Analytics

Big Data erleichtert die Mobilitätsforschung und -kontrolle. Wo Personenströme von Bedeutung sind, wird verlässliche Planung zur Herausforderung. Die Auslastung des Nahverkehrs und die Positionierung von Werbung ließen sich bisher nur mit aufwendigen Handzählungen und rückblickend beantworten. Mit Big Data in Verbindung mit Lokationsdaten 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.

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. 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.

2.11 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, Dienstanbieter von optimierten Erträgen aus der Bereitstellung ihrer Services. Die Erfassung, Aggregation und Aussteuerung von Daten und Werbeeinheiten 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.

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 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.

2.12 Industrie 4.0

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

2.13 Telekommunikation

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 Telekom-

unikationsinfrastruktur 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. Dynamische Analytics-Lösungen nutzen beispielsweise die Fähigkeit von Telekommunikationsunternehmen, Bewegungsmuster zu erstellen, die auf Lokationsdaten der Teilnehmer sowie auf CRM³-Daten basieren und mittels statistischer Methoden Bewegungsströme darstellen können. Verkehrsbetreiber erhalten so wichtige Informationen für die Bedarfsplanung und -steuerung.

2.14 Verbesserung des Datenschutzes in Rechenzentren

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

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

3. 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.

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 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. 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 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.

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 morbiditätsorientierten Risikostrukturausgleichs⁴ anfallen, funktionierende von nicht funktionierenden Behandlungsmustern getrennt werden. Zumindest ließe sich mit einer signifikanten Wahrscheinlichkeit sagen, welche Behandlung „normalerweise“ Erfolg verspricht und welche „fast nie“ wirkt. Sicher wären alle Patienten dankbar, von den Erfahrungen „aller“ anderen vergleichbaren Fälle zu profitieren.

4 Risikostrukturausgleich der gesetzlichen Krankenkassen

Hinderungsgründe

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

- 1.** Zweckbindung im Bundesdatenschutzgesetz: Der Reiz und der Mehrwert vieler Big-Data-Projekte entsteht im explorativen Erforschen und Kombinieren bisher getrennt erfasster Einzeldatenbestände. Damit lassen sich beispielsweise Fragestellungen nach dem Zusammenhang vom Wetter (Temperaturdaten) und dem 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. 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.
- 2.** Starke Anforderungen an Anonymisierung: Der übliche und häufig verfolgte Ausweg bei nicht ausreichenden Einwilligungen ist die 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.

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.
- Auch fehlt es an erprobten Technologien, die diese Anonymisierung und Pseudonymisierung so einfach machen, dass es für alle Beteiligten transparent und handhabbar bleibt. Erfolgversprechende Ansätze beziehen auch Hardwarekomponenten mit ein oder arbeiten mit ausgefeilten Zertifizierungshierarchien.
- 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.

4. 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 Start-ups. 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: 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.

- Internet der Dinge und von Software bestimmte Produkte: Mit dem Internet der Dinge beginnt eine engmaschige Verzahnung von analoger und digitaler Welt. Angefeuert 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 -design von Haushaltsgeräten, Konsumelektronik, Kleidung oder auch Autos. Der Wert eines Produktes aus Verbrauchersicht wird immer stärker durch intelligente Software-Funktionalität und kontextsensitive Bereitstellung von Daten geprägt. Die Einzigartigkeit eines Produkts ist durch Software definiert; die Bereitstellung und Verarbeitung von produkt- und kontextrelevanten Daten spielt für den Anwender eine zunehmende Rolle. Sie bestimmt somit maßgeblich den Differenzierungsfaktor eines Produktes sowie die Zahlungsbereitschaft der Verbraucher.
- Software-bestimmtes 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. 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.

Diese drei beschriebenen Technologie- und Strategietrends verstärken sich wechselseitig.

In dem genannten Zeitraum von drei bis fünf Jahren 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 FinTech⁵-Start-ups mit ihren anwenderfreundlichen und mobilen Lösungen in die Finanzströme einklinken und den etablierten Finanzkonzernen die jungen und zahlungsfreudigen Verbraucher abnehmen.

5 <https://de.wikipedia.org/wiki/Finanztechnologie>

5. Bedeutung der Datenwirtschaft in der zukünftigen Wirtschaft

Der Einsatz von Big-Data-Technologien ist Gegenstand eines intensiven öffentlichen Diskurses. Im Kapitel 5 werden zehn ausgewählte Fragenkomplexe aufgegriffen.

5.1 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. Aus Sicht des Datenschutzes ist in diesen Fällen von besonderem Interesse, welche Regeln zur Anwendung kommen. In der Europäischen Union (EU) ist das auf der Richtlinie 95/46 EG⁶ 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 Auslegungstreitigkeiten sowohl für den Datenverkehr innerhalb der EU als auch für den Datentransfer in Drittstaaten, also außerhalb der EU und des Europäischen Wirtschaftsraums (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.

⁶ [https://de.wikipedia.org/wiki/Richtlinie_95/46/EG_\(Datenschutzrichtlinie\)](https://de.wikipedia.org/wiki/Richtlinie_95/46/EG_(Datenschutzrichtlinie))

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-Verarbeitungsprozessen Beteiligten, egal, ob diese Daten innerhalb oder außerhalb der Europäischen Union verarbeiten.

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.

5.2 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.

Kommerzielle Big-Data-Projekte werden in der Öffentlichkeit oft als intransparent wahrgenommen oder dargestellt. 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.

5.3 Transparenz bei Big-Data-Analysen

Transparenz bei Big-Data-Analysen und -Anwendungen 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.

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, 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. 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.

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.

5.4 Wie Analyseergebnisse zustande kommen

In immer mehr Organisationen werten eigens dafür abgestellte Experten die Datenbestände systematisch aus. Diese „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.

5.5 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. Die Medien rücken unter dem Schlagwort Big Data teilweise Gefahren und negative Beispielfälle in den Vordergrund („Big Brother“). Die Verbraucher verlangen, berechtigter Weise, 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.

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. Vor diesem Hintergrund ist es für Unternehmen besonders wichtig, gegenüber ihren Verbrauchern transparent über Big-Data-Lösungen zu informieren.

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. 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 – 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 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 BDSG⁷, 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.

7 <https://dejure.org/gesetze/BDSG>

Bedeutung von Leitlinien

Big-Data-Analysen bergen also vielfältige Herausforderungen, nicht zuletzt technischer und wirtschaftlicher Natur. Die im Kapitel 6 vorgestellten Leitlinien sollen den Unternehmen eine Richtschnur für einen ethisch und juristisch einwandfreien Einsatz von Big-Data-Analysen aufzeigen.

Vorgestellt 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.

5.6 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.

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.

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 bieten.

5.7 Datensammlung im Automatismus

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.

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.

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.

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.

5.8 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.

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-Schönberger⁸. 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.

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 aufgegriffen, wobei keine festen Löschrufen vorgesehen wurden und das Recht vornehmlich einer Ausweitung des Rechts auf informationelle Selbstbestimmung gleichkam. Aufgrund zahlreicher Kritik wurde das Recht im laufenden Gesetzgebungsverfahren vom Europäischen Parlament auf ein Recht auf Löschung beschränkt.

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

8 https://de.wikipedia.org/wiki/Recht_auf_Vergessenwerden

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.

2014 entschied der Europäische Gerichtshof (EuGH), dass es für den Einzelnen möglich sein muss, die Löschung von Links auf personenbezogene Daten zu veranlassen. Diese Entscheidung wurde so umgesetzt, dass die betreffenden URLs⁹ 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. Jedoch stellt die Limitierung trotzdem eine massive Einschränkung der informationellen Selbstbestimmung dar. Umgekehrt würde ein absolutes Löschrrecht 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 Löschen utopisch. Vielmehr ist ein 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 darum,

über die Möglichkeit zu selbstbestimmtem Löschen zu 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 „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.

5.9 Umgang mit Wahrscheinlichkeiten und Verhaltensprognosen

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.

⁹ https://de.wikipedia.org/wiki/Uniform_Resource_Locator

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. Bestimmte Analysen müssen ausgeschlossen bleiben, beispielsweise genetische Analysen für berufliche Zwecke. Auch dürfen staatliche Zwangsmaßnahmen nicht mit einer bloßen Wahrscheinlichkeitsprognose begründet werden.

5.10 Erkenntnisse aus der Kommunikation in sozialen Netzwerken

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. Unternehmen sollten hier umsichtig agieren und die Interessen der Betroffenen angemessen berücksichtigen. 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. Oftmals geht die Anonymisierung der Daten dabei nicht mit einem Erkenntnisverlust einher.

6. 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 berechnete 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 Datennutzen 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.

7. Big Data und Datenschutz: Politischer Handlungsbedarf

Dieses Positionspapier 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.

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.

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?
- 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. 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 für die Selbstbestimmung oder Freiheit des Einzelnen zu ignorieren.

Chancen und Risiken müssen stets in einen angemessenen Ausgleich gebracht werden; 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 Grundrechtswirklichkeiten – ausblendet.

Angesichts der globalen Nutzung von Big-Data-Produkten sind dabei auch die unterschiedlichen kulturellen Sichtweisen zu berücksichtigen. Letztlich erstrebenswert wäre ein globaler Datenschutz.

EU-Datenschutz-Grundverordnung weiterentwickeln

Zum Start der Verhandlungen über die Details der EU-Datenschutz-Grundverordnung forderte der Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V. (bitkom) Nachbesserungen, 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 wird unterstrichen, dass eine einseitige Verengung der öffentlichen Diskussion auf Wirtschaft contra Bürgerrechte an den Realitäten vorbei geht.

Die Projektgruppe Smart Data 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.

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 Verkehrlenkung. Anonymisierte, also irreversible de-personalisierte Daten fallen aus dem Anwendungsbereich der Verordnung heraus, pseudonymisierte, also reversible de-personalisierte Daten dagegen nicht. Sie werden rechtlich behandelt Klardaten. Gerade die pseudonymisierten Daten 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 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 Ratsfassung¹⁰ 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 <http://www.cr-online.de/26378.htm>

8. Autorenverzeichnis

Prof. Dr. Christoph Bauer
ePrivacy GmbH

Andreas Hufenstuhl
CSC Deutschland GmbH

Dr. Jens Schefzig
Osborne Clarke

Martin Birkmeier
FIR e.V. an der RWTH Aachen

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

Barbara Schmitz
Kabel Deutschland Vertrieb und
Service GmbH

Arnd Böken
Graf von Westphalen
Rechtsanwälte Partnerschaft

Dr. Thomas Keil
SAS Institute GmbH

Markus Schröder
Tembit Software GmbH

Dr. Mikio L. Braun,
Technische Universität Berlin

Ralf Konrad
T-Systems International GmbH

Maximilian Thess
Lemma Group UG

Patrick von Braunmühl
Selbstregulierung Informationswirtschaft e.V.

Holm Landrock
Experton Group AG

Dr. Carlo Velten
Crisp Research AG

Florian Buschbacher
PricewaterhouseCoopers AG
Wirtschaftsprüfungsgesellschaft

Dr. Mario Lenz
empolis GmbH

Dr. Mathias Weber
Bitkom e.V.

Martin Buske
DataXu GmbH

Dr. Ralf Meinberg
Deutsche Telekom AG

Alexander Widak
Fiducia & GAD IT AG

Susanne Dehmel
Bitkom e.V.

Dr. Flemming Moos
Osborne Clarke

Dr. Frank Wisselink
Detecon International GmbH

Guido Falkenberg
Software AG

Steffen Nienke
FIR e.V. an der RWTH Aachen

Dr. Andreas Zolper
Deutsche Telekom AG

Dr. Jörg Friedrichs
Deutsche Telekom AG

Swen Obermann
Capgemini Deutschland GmbH

9. Mitwirkende Projektgruppe Smart Data

Leitung

Prof. Dr. Christoph Meinel
Hasso-Plattner-Institut

Dr. Norbert Koppenhagen
SAP SE

Reiner Bildmayer
SAP SE

Dr. Karina Lott
RELX Group

Stefan Vaillant
Cumulocity GmbH

Guido Falkenberg
Software AG

Dr. Pablo Mentzins
Bundesverband Informationswirtschaft,
Telekommunikation und neue Medien e.V.

Manuela Wagner
Karlsruher Institut für Technologie

Helmut Greger
Charité – Universitätsmedizin Berlin

Martin Peuker
Charité – Universitätsmedizin Berlin

Mathias Weber
Bundesverband Informationswirtschaft,
Telekommunikation und neue Medien e.V.

Nina Hrkalovic
Gesellschaft für Informatik e.V.

Alexander Rabe
Gesellschaft für Informatik e.V.

Markus Widmer
Intel GmbH

Prof. Dr. Michael Laskowski
RWE Deutschland AG

Hannes Schwaderer
Intel GmbH

Dr. Alexander Lenk
FZI Forschungszentrum Informatik

David Schwalb
Hasso-Plattner-Institut für
Softwaresystemtechnik GmbH

Sven Löffler
T-Systems International GmbH

Günther Stürner
ORACLE Deutschland B.V. & Co. KG



Leitlinien für den Big-Data-Einsatz im Überblick

Chancen und Verantwortung

Ausgabe: Oktober 2015

Herausgeber

Plattform „Innovative Digitalisierung der Wirtschaft“ im Nationalen IT-Gipfel
Fokusgruppe Intelligente Vernetzung / Projektgruppe Smart Data

Ansprechpartner

Prof. Dr. Christoph Meinel
Hasso Plattner Institut
christoph.meinel@hpi.de

Dr. Norbert Koppenhagen
SAP SE
norbert.koppenhagen@sap.com