



BIG DATA

Big Data

Big Data

Eine Untersuchung des iRights.Lab im Auftrag des
Deutschen Instituts für Vertrauen und Sicherheit im Internet (DIVSI)



Hamburg, Januar 2016

Deutsches Institut für Vertrauen und Sicherheit im Internet (DIVSI)

Mittelweg 110 B,
20149 Hamburg
www.divsi.de

Matthias Kammer, Direktor

Joanna Schmölz,

Wissenschaftliche Leitung

Afia Asafu-Adjei, Projekte

Dr. Dirk Graudenz,

Projektleitung

iRights.Lab

Almstadtstr. 9/11,

10119 Berlin

www.iriights-lab.de

Philipp Otto, Projektleitung

Dr. Till Kreutzer, Autor

Henning Lahmann, Autor

Jan Schallaböck, Autor

Unter Mitwirkung von: Valie Djordjevic, Wiebke Glässer,

Eike Gräf, Tom Hirche, Hanka Holzapfel, Jana Maire,

Julia Schrader



Creative-Commons-Lizenz: CC BY-ND 3.0 DE

Die Texte dieses Werks sind unter der Creative-Commons-Lizenz vom Typ „Namensnennung – Keine Bearbeitung 3.0 Deutschland“ lizenziert. Um eine Kopie dieser Lizenz einzusehen, besuchen Sie <http://creativecommons.org/licenses/by-nd/3.0/de>. Diese Lizenz beinhaltet unter anderem, dass die Texte bei Nennung des/der Autoren und dieser Publikation als Quelle ohne Veränderung veröffentlicht und weitergegeben werden dürfen. Ausgenommen von dieser Lizenz sind alle Nicht-Text-Inhalte wie Fotos, Grafiken und Logos.

INHALTSVERZEICHNIS

Vorwort Matthias Kammer	6	2.3 Anwendungsfall Autoversicherung	70
Vorwort Phillipp Otto	8	2.4 Anwendungsfall Verkehrsleitsysteme	76
Zusammenfassung	10		
1. Einführung in die Thematik	13	3. Smart Health: Tracking und Wearables	80
1.1 Big Data	14	3.1 Begriffsdefinition	82
Andrea Voßhoff: „Big Data – Chancen und Risiken“	18	Karl Max Einhüpl : „Big Data – Chancen und Risiken im Gesundheitsbereich“	84
1.2 Tracking	24	3.2 Smart Health als Anwendungsfeld für Tracking	89
1.3 Entwicklung und Stand der Technik	26	Öffentliche Veranstaltung in Berlin: „Treuer Assistent oder Trojaner am Körper? – Wie Gesundheitstracking unseren Alltag verändert“	90
Silke Jandt: „Rechtskonforme Technik wäre ein großer Vorteil für alle“	30	3.3 Wearables als Tracking-Technologie	94
1.4 Anwendungsgebiete	35		
Sabrina Juran: „Big Data für Armutsbekämpfung und nachhaltige Entwicklung“	38	4. Mögliche Kodifizierungen im Bereich Big Data	110
1.5 Big Data im Gesamtprojekt „Braucht Deutschland einen Digitalen Kodex“?	43	4.1 Übergreifende Aspekte der Anwendungsfelder	110
1.6 Der Projektverlauf	44	4.2 Grundkonflikte	111
Die Expertengruppe	46	Tobias Keber: „„Privacy made in Germany“ könnte ein Verkaufsargument sein“	113
1.7 Aufbau des Dokuments	53	4.3 Algorithmenethik	119
2. Smart Mobility: Tracking im Bereich Mobilität	54	5. Epilog	120
2.1 Begriffsdefinition	56	Annex	121
Marit Hansen: „Es besteht eine Lücke zwischen dem Datenschutz und der Technik“	58	Analyse des Themenkomplexes „Big Data“	121
2.2 Der Einsatz von Tracking für Smart Mobility-Anwendungen	64	Quellen und Literaturhinweise	137
Öffentliche Veranstaltung in Stuttgart: „Vernetzte Mobilität – Erweiterung der persönlichen Autonomie oder Eingrenzung der Privatsphäre?“	66	Über die an diesem Bericht beteiligten Organisationen und Institutionen	139

Big Data

Es ist nicht einfach, sich Big Data und allem, was damit zusammenhängt, halbwegs wertfrei zu nähern. Dabei bezeichnet der Begriff in seinem Ursprung nur die Verarbeitung von großen, komplexen und sich schnell ändernden Datenmengen. Im Alltagsgebrauch ist er jedoch zu einem Buzzword verkommen und wird derzeit meist mit negativem Anklang verwendet.

So erfolgen häufig Hinweise auf Gefahren von Big Data, wenn es um die zunehmende Überwachung der Menschen durch Vorratsdatenspeicherung, die Verletzung der Persönlichkeitsrechte von Kunden durch Unternehmen, die steigende Intransparenz der Datenspeicherung in Clouds oder auch um unerwünschte datenbasierte Werbung geht, die bei der Nutzung von Internet und Handy anfallen.

Big Data deckt also auch Bereiche ab, die bisher als privat galten. Der Wunsch der Wirtschaft und von Behörden, möglichst umfassenden Zugriff auf diese Daten zu erhalten und die gewonnenen Erkenntnisse zu nutzen, gerät zunehmend in Konflikt mit Persönlichkeitsrechten des Individuums.

Ist Big Data also komplett negativ?

Auf jeden Fall produziert der Begriff für das deutsche Gemüt aktuell vor allem Ängste und ver-

stellt dadurch den Blick auf die positiven Seiten. Er wirkt eher aggressiv.

Was steckt eigentlich hinter dem Begriff „Big Data“? In seinem aktuellen Gebrauch meint er die Verarbeitung großer Datenmengen, um bisher verborgene Zusammenhänge sichtbar und nutzbar zu machen. Eine Verbindung also, die viele Nutzer mit Sorge betrachten. Doch haben wir überhaupt eine andere Chance? Laut Berechnungen von IT-Marktforschern verdoppelt sich das weltweite Datenvolumen alle zwei Jahre. Bis 2020 soll es auf rund 40 Zettabytes (eine Zahl mit 21 Nullen) anwachsen. Es liegt auf der Hand, dass ein solch gigantischer Berg an Informationen unseres Internet-Zeitalters längst zu komplex ist, um mit klassischen Methoden der Datenverarbeitung oder gar händisch ausgewertet werden zu können.

Big Data ist also unaufhaltsam. Denn die Digitalisierung hat längst alle Lebensbereiche erreicht. Sie ist dabei – und das bestreitet kaum noch jemand – unsere Lebensumstände insgesamt umzukrempeln. Deshalb macht es durchaus Sinn, wenn wir die Erkenntnisgewinne durch Datenanalysen nutzen und versuchen, daraus Positives abzuleiten. Das Motto dieses Umdenkens könnte etwa lauten: Mehr Daten, bessere Erkenntnisse.

Foto: Frederike Heim



Matthias Kammer,
Direktor
DIVSI – Deutsches Institut für Vertrauen
und Sicherheit im Internet

Nach meiner Einschätzung gehen Big Data und die damit verbundenen Chancen nicht in Einklang mit dem Paradigma der Datensparsamkeit, wie es sich aus dem geltenden Datenschutzrecht ergibt. Nutzer begrenzen im Alltag in der Regel auch nicht den Umfang ihrer persönlichen Daten. Also sind wir doch besser beraten, darüber nachzudenken, wie sich die Chancen von Big Data verstärkt fördern lassen. Zugleich sollte Missbrauch aber schärfer geahndet und auch geächtet werden, als es die Instrumente bislang zulassen. Ich meine, dass für Unternehmen Beschädigungen an ihrer Reputation weit stärker zählen als Ordnungsgelder.

Wenn wir die Chancen verstärkt zulassen, wird die sorgfältige Auswertung von Daten uns beispielsweise für Mobilität oder Gesundheit neue Chancen eröffnen. Im Hinblick auf Smart Health hat das gerade eine weitere Info-Veranstaltung von DIVSI gezeigt, die im Rahmen des Projektes „Braucht Deutschland einen Digitalen Kodex?“ in Berlin eine Vielzahl von Experten im Meistersaal versammelt hatte. Gesundheitstracking wird demnach kaum zu stoppen sein. Big Data ist dabei der Schlüssel, der dem einzelnen Nutzer aus den gewonnenen Fakten heraus wertvolle Hinweise liefern kann.

Natürlich darf man die Risiken nicht vergessen: Was geschieht mit den so umfangreich gesammelten Daten? Unsere Gesellschaft muss intensiv darüber nachdenken

und Regeln dafür finden, wie sich Missbrauch verhindern lässt. Dabei sollten wir Chancen und Risiken nicht getrennt diskutieren. Die – zugegeben komplexe – Aufgabe liegt darin, beide Stränge miteinander zu verknüpfen. Gelingt dies, wird sich dadurch das Vertrauen der Menschen in das Internet festigen.

DIVSI hat 2013 das Projekt „Braucht Deutschland einen Digitalen Kodex?“ mit dem Berliner Think Tank iRights.Lab in Angriff genommen. Die Grundfrage ist längst positiv beantwortet: Ein Digitaler Kodex kann die Lücke zwischen sozialen Normen der analogen Welt und den neuen Herausforderungen der digitalen Welt schließen.

In Phase zwei des Projekts haben wir uns seit Herbst 2014 mit Big Data befasst, zugespielt auf die Schwerpunkte Smart Health und Smart Mobility. Dieser Arbeitsabschnitt analysiert die Auswirkungen der neuen Technologien und zeigt Wege auf, mit den Herausforderungen umzugehen.

Die Erkenntnisse aus vielen Expertenrunden fasst dieser Bericht zusammen. Er zeigt vor allem auch, dass im Kontext von Big Data eine umfassende gesellschaftliche Debatte zwingend notwendig ist. Ich bin überzeugt, dass die hier präsentierten Erkenntnisse gedankliche Anstöße in dieser Richtung leisten können.

Big Data: Die Revolution gestalten!

„Lentement, depuis vingt ans, le surnaturel est sorti de nos âmes. Il s'est évaporé comme s'évapore un parfum quand la bouteille est débouchée.“¹

Guy de Maupassant, „Le Fantastique“, Le Gaulois, 7. Oktober 1883

Im 19. Jahrhundert vertrieb die Wissenschaft den Sinn für das Übersinnliche aus den Seelen der Menschen. Aberglaube und Irrationales wurden durch Fortschrittsdenken und Wissenschaftsglauben abgelöst. Im 21. Jahrhundert bringt die Wissenschaft ein neues Gefühl des Mysteriösen und Unerklärlichen in unsere Gesellschaften zurück: Die teilweise noch undefinierten Möglichkeiten und Funktionsweisen von Big-Data-Anwendungen und mächtiger Algorithmen erobern unseren Alltag und hinterlassen den Eindruck, unbekanntem Mächten ausgeliefert zu sein. So wie der französische Erzähler Guy de Maupassant 1883 versuchte, die damaligen Entwicklungen in Worte zu fassen, so sitzen wir heute vor technischen Entwicklungen, die für uns nur

schwer fassbar scheinen, und versuchen, sie zu interpretieren und eine gesellschaftliche Antwort zu finden.

Wo sind die roten Linien beim Tracking unserer physischen Bewegungen, unseres Pulsschlags und unseres Nutzungsverhaltens, wenn wir uns in der digitalen Welt bewegen? In welchem Moment sind die allgegenwärtigen Unterstützungssysteme nicht mehr die helfende Hand in unserem Leben, sondern die hässliche Pranke, die uns eine Ohrfeige verpasst? Was wollen wir wissen und was besser nicht, weil es uns überfordert?

Die zunehmend flächendeckende digitale Aufzeichnung unseres Lebens braucht gesellschaft-

¹ „Langsam, seit zwanzig Jahren, ist das Übernatürliche aus unseren Seelen verschwunden. Es hat sich verflüchtigt, wie ein Parfum sich verflüchtigt, wenn die Flasche geöffnet ist.“ Siehe auch http://lettres.ac-rouen.fr/archives_bac/afrique/r019-mal.htm.



Foto: Bettina Volke

Philipp Otto,

Leiter des Projekts „Braucht Deutschland einen Digitalen Kodex?“ und Gründer des unabhängigen Think Tanks iRights.Lab

lich definierte Regeln. Immer, wenn wir nicht mehr weiterwissen, rufen wir nach Regulierung und dem Gesetzgeber. Doch in welcher Situation und in welcher Konstellation brauchen wir ein Gesetz – einen sanktionierenden Eingriff in eine technische Entwicklung, von der wir heute oft noch nicht wissen, wohin sie unsere Gesellschaft und uns ganz persönlich führen wird? Wir befinden uns in einer technisch definierten digitalen Kulturrevolution – die Entwicklungen und Diskussionen der Zukunft werden zeigen, wie wir damit umgehen werden.

In dieser Publikation finden Sie Hintergrundinformationen, Einschätzungen, Bewertungen und Analysen zu zwei hochrelevanten, zunehmend

digital durchdrungenen Bereichen unseres Alltags: Mobilität („Smart Mobility“) und Gesundheit („Smart Health“). In beiden Bereichen liefern wir Ihnen in diesem Bericht Material für eine gesellschaftliche Debatte, die überfällig ist. Ziel der Untersuchung ist eine Zuspitzung auf die neuralgischen Punkte, die wir gesellschaftlich klären müssen. Nehmen Sie sich Zeit, reflektieren Sie bei der Lektüre Ihren Alltag und versuchen Sie zu definieren, welche ethischen Prämissen Sie bei Bewertungen digitaler Innovationen zugrunde legen. Machen Sie sich dabei aber klar, dass wir technische Entwicklungen nicht aufhalten können.

Ich bin gespannt zu erfahren, zu welchen Erkenntnissen Sie kommen werden!

Zusammenfassung

Der vorliegende Bericht beschreibt die Chancen und Risiken des Einsatzes von Big Data auf den Gebieten Smart Mobility und Smart Health und identifiziert die daraus resultierenden Konfliktlinien und gesellschaftlichen Herausforderungen. Das dem Bericht zugrunde liegende Projekt ist Teil des Gesamtvorhabens „Braucht Deutschland einen Digitalen Kodex?“, das 2013 vom Deutschen Institut für Vertrauen und Sicherheit im Internet (DIVSI) ins Leben gerufen wurde. Ziel der im Herbst 2014 gestarteten zweiten Projektphase ist es, zu untersuchen, ob das Modell eines Digitalen Kodex auch in der konkreten Anwendung trägt und sich realisieren lässt.

Der Projektteil „Big Data“ befindet sich in einem Prozess, der noch nicht abgeschlossen ist. In einem ersten Schritt auf dem Weg zu einem möglichen Digitalen Kodex für den Umgang mit Big Data wurden in

der hier dokumentierten Projektphase zunächst konkrete Konfliktfelder beschrieben, die durch Big Data entstehen. Da das Phänomen äußerst facettenreich ist, wurde die Thematik auf zwei Anwendungsgebiete eingegrenzt: Smart Mobility und Smart Health. An ihnen wurden exemplarisch die Chancen und Herausforderungen von Big Data herausgearbeitet, um Regelungsbedürfnisse zu identifizieren, denen dann mit einem zu schaffenden Digitalen Kodex begegnet werden könnte. Ein Projektschwerpunkt war Tracking als eine der primären Methoden, bei der die in Big-Data-Anwendungen eingesetzten großen Datenmengen anfallen.

Bei Big-Data-Anwendungen werden computergestützt große Datenmengen analysiert und weiterverwendet. Diese sind insbesondere durch die immer weiter zunehmende Digitalisierung vieler Lebensbereiche in

den vergangenen Jahren entstanden, wobei Tracking-Technologien eine der Hauptquellen für die Erzeugung dieser Daten sind. Tracking kommt in immer mehr Anwendungen zum Einsatz, so zum Beispiel in Smartphones, Fitnessarmbändern oder in Blackboxes, die in Autos eingebaut werden. Die Geräte zeichnen fortlaufend bestimmte Werte auf und übermitteln die so gewonnenen Daten an Server der Anbieter, wo sie gespeichert werden und für Big-Data-Auswertungen zur Verfügung stehen.

Im Bereich Smart Mobility sind telematikbasierte KFZ-Ver sicherungen ein aktuelles Anwendungsfeld. Tarife werden dabei unter anderem auf Basis des tatsächlichen Fahrverhaltens des Versicherungsnehmers angepasst. Zu diesem Zweck wird zumeist eine Blackbox in das Auto des Kunden eingebaut, die das Fahrverhalten des Fah-

ers verfolgt und aufzeichnet. Auf Basis einer Datenanalyse werden schließlich die Prämien errechnet und entsprechend angepasst.

Diese Art von Versicherungsmodell verspricht eine Reihe von Chancen. So werden die Autoversicherer in die Lage versetzt, profitablere Versicherungsmodelle anbieten zu können. Zugleich können die Autofahrer profitieren, insbesondere jene, die bislang aufgrund ihrer Zuordnung zu einer bestimmten Gruppe – wie beispielsweise Fahranfänger – ungeachtet ihres individuellen Fahrstils hohe Prämien zahlen mussten. Darüber hinaus könnte sich die Sicherheit auf den Straßen erhöhen und die Umweltbelastung sinken, worin ein gesamtgesellschaftlicher Nutzen zu sehen wäre.

Zugleich birgt ein solches Modell Risiken. So ermöglicht das Tracking, umfassende Bewegungsprofile von Autofahrern zu erstellen. Darüber hinaus ist an Personengruppen zu denken, die beispielsweise im Schichtdienst arbeiten und deshalb ihr Fahrzeug zu Zeiten nutzen müssen, die als risikobehaftet identifiziert werden, und deshalb höhere Tarife zahlen müssen.

Ein weiteres Beispiel für eine Anwendung im Bereich Smart Mobility ist die datengetriebene Verkehrslenkung in Ballungsräumen, die auf Grundlage von Bewegungs- und Geolokationsdaten die Verkehrsströme zu optimieren versucht. Dies kann den Verkehrsteilnehmern helfen, Zeit einzusparen. Durch die

daraus folgende Ressourcenoptimierung werden zudem weniger Emissionen freigesetzt. Den kommunalen Planungsbehörden bieten sich darüber hinaus Möglichkeiten, Infrastrukturprojekte gezielter zu planen und durchzuführen, während die Verkehrsbetriebe die ihnen zur Verfügung stehenden Mittel präziser einsetzen können. Demgegenüber kann das Tracking der Bewegungen der Bürger im städtischen Raum trotz anonymisierter Daten zum Gefühl beitragen, unter Überwachung zu stehen, was sich verhaltensändernd auswirken kann.

Auch im Gesundheitssektor wird im Bereich Smart Health immer mehr auf Big-Data-Anwendungen zurückgegriffen. Die dafür notwendigen Datenmengen werden unter anderem durch Wearables erzeugt, also beispielsweise durch Fitness-Armbänder oder Smart Watches, die Vitalwerte ihres Trägers aufzeichnen.

Gerade im Gesundheitssektor sind mit Big Data große Hoffnungen verbunden, die Prävention, Diagnose und Therapie von Krankheiten zu verbessern. So können Ärzte den Gesundheitszustand ihrer Patienten präziser überwachen; Krankenhäuser können aus großen Datensätzen von Patienten Erkenntnisse gewinnen, anhand derer sie die verfügbaren Ressourcen besser verteilen können; auch zur Vorhersage von Epidemien kann Big Data nützlich sein. In besonderem Maße verspricht sich die medizinische Forschung Chancen durch Big Data, da es mit-

tels Wearables künftig leichter sein dürfte, Probanden für Studien zu gewinnen. Dadurch entsteht leichter eine ausreichend große Datenbasis für medizinische Erkenntnisse. Auch das Krankenversicherungssystem kann durch die Informationen, die durch Wearables gewonnen werden, neue Tarifmodelle schaffen, die Kosten einsparen helfen und Ressourcen besser nutzbar machen. Nicht zuletzt profitieren die Nutzer solcher Geräte selbst von den genannten Möglichkeiten.

Auch der Einsatz von Big Data im Gesundheitsbereich birgt Risiken. So vergrößert sich für den Nutzer die Gefahr der Fremdbestimmtheit, wenn durch die konstante Überprüfbarkeit und Vergleichbarkeit von Gesundheitsdaten gesellschaftlich neu definiert wird, welche körperlichen Zustände als gesund, also „richtig“, und welche als krank, also „falsch“, gelten. Diese Neudefinition kann zu einem erhöhten Anpassungsdruck führen. Eine Gefahr besteht zudem darin, dass die Möglichkeit, sich gegen die Nutzung von Wearables zu entscheiden, an gesellschaftlicher Akzeptanz verliert. Da es sich bei Gesundheitsdaten um besonders persönliche und sensible Daten handelt, stellt sich außerdem die Frage, wie Datenschutz und Datensicherheit gewährleistet werden können, um Missbrauch im Umgang mit den Daten zu verhindern. Für die medizinische Forschung schließlich besteht zumindest potenziell das Problem, dass medizinische Studien, die mit-

tels Wearables durchgeführt werden, für das Phänomen der Stichprobenverzerrung besonders anfällig sein könnten, wenn sich zum Beispiel die Studienteilnehmer ausschließlich aus solchen Bevölkerungsteilen rekrutieren, die sich ein solches Gerät leisten können.

Dieser Bericht beginnt mit der Darstellung von insgesamt drei Beispielen der beiden Anwendungsgebiete Smart Mobility und Smart Health, gefolgt von der Herausarbeitung der Chancen, Risiken und Konfliktlinien und schließlich der Identifizierung der Herausforderungen für die Gesellschaft im Umgang mit Big Data. Letztere könnten in einer weiteren Projektphase als Grundlage für einen Digitalen Kodex dienen, der mit den relevanten Akteuren auszuhandeln wäre. Den Beispielen ist gemein, dass die Datenerhebungen oftmals mittels Tracking-Technologien erfolgen und stets mit einem Eingriff in die Privatsphäre verbunden sind, insofern, dass die Auswer-

tung der Daten Rückschlüsse über das Leben des Einzelnen zulässt. Zu vertiefende Themen sind:

- Big Data stellt die Grundprinzipien des traditionellen Datenschutzrechts wie Datensparsamkeit und Zweckbindung infrage. Es muss daher geklärt werden, inwieweit ein Digitaler Kodex einen Beitrag dazu leisten kann, speziell auf Big Data abgestimmte Lösungen für die dadurch entstehende Problematik zu finden.
- Wie kann gewährleistet werden, dass die Nutzung von Tracking-Geräten freiwillig bleibt und eine Nichtteilnahme nicht mit unangemessen großen Nachteilen verbunden ist, wenn Tracking-Technologien und darauf basierende Big-Data-Anwendungen immer weitere Verbreitung finden? Hat dieses Prinzip der Freiwilligkeit überhaupt weiterhin Bestand?
- Weiterhin ist klärungsbefähig, wie die Neutralität

und Transparenz jener Algorithmen gewährleistet werden kann, die Big-Data-Anwendungen zugrunde liegen. Diese Aspekte lassen sich unter dem Begriff der Algorithmenethik zusammenfassen. Wie und von wem können Algorithmen daraufhin überprüft werden, ob sie beispielsweise bestimmte Bevölkerungsgruppen unangemessen benachteiligen? Welche Mechanismen müssten geschaffen werden, um die Verwendung solcher diskriminierender Algorithmen zu unterbinden?

Das Projekt hat bisher gezeigt, dass in Bezug auf Tracking und Big Data eine Reihe von Konfliktlinien existieren, die den Bedarf nach einer gesellschaftlichen Debatte aufwerfen. Eine solche könnte ergeben, dass für Big Data neue oder angepasste Regeln benötigt werden, für die sich alternative Regulierungsansätze wie ein Digitaler Kodex eignen könnten.

1. Einführung in die Thematik

Das im Jahr 2013 gestartete Projekt „Braucht Deutschland einen Digitalen Kodex?“ des Deutschen Instituts für Vertrauen und Sicherheit im Internet (DIVSI) ging im Herbst 2014 in die zweite Phase. In der ersten Phase des Projekts (April 2013 bis Mai 2014) wurden zunächst grundlegende Aspekte behandelt. Hier wurde untersucht, ob sich das Modell eines Digitalen Kodex prinzipiell dazu eignet, bestehende Lücken herkömmlicher Regulierung oder sozialer Normung zu schließen, die in der digitalen Welt entstanden sind. Zu diesem Zweck wurden zunächst die Unterschiede zwischen der physischen und der digitalen Welt herausgearbeitet, die für Regulierungsaspekte relevant sind, Regelungslücken identifiziert und grundlegende Aspekte der Verantwortung im Netz untersucht. Zudem wurde definiert, was einen Digitalen Kodex ausmacht. Das Ergebnis war, dass unter einem Digitalen Kodex eine alternative Form der Regulierung zu verstehen ist, die Elemente klassischer Regelkonzepte, wie Gesetze, Selbstverpflichtungen und soziale Normen, kombiniert. Ein wesentliches Merkmal eines Digitalen Kodex ist ein Diskurs mit den relevanten Akteuren, in dem der Kodex erarbeitet, ein- und umgesetzt wird. Der Einsatz eines solchen prozessorientierten Verfahrens bietet sich besonders bei auf das Netz bezogenen Regelwerken an. Durch diese Form der Erarbeitung und Umsetzung – so der Ansatz – lassen sich Defizite klassischer Regulierungsformen bei Internet-Sachverhalten beheben, wie etwa mangelnde Akzeptanz und Durchsetzung. Im Ergebnis kam die erste Projektphase zu dem Schluss,

dass ein Digitaler Kodex ein vielversprechendes Regulierungsmodell darstellt, das weiter untersucht werden sollte.

In der zweiten, nun abgeschlossenen, Projektphase wurde der Ansatz weiter vertieft. Ziel war es, anhand von konkreten Anwendungsfeldern für einen solchen Digitalen Kodex zu untersuchen, ob das Modell in der Anwendung trägt und sich realisieren lässt. Die Prüfung der Praxistauglichkeit kann – auch dies war ein Ergebnis der ersten Projektphase – nur anhand spezifischer Anwendungsfelder beantwortet werden.

Für die zweite Phase des Projekts wurden beispielhafte Testfelder ausgewählt, anhand derer untersucht wurde, welche konkreten Schritte zur Umsetzung nötig sind. Bei dem in diesem Bericht dokumentierten Teilprojekt geht es um Big Data. Das Thema wurde dahingehend untersucht, ob die Konzeption und Aushandlung eines spezifischen Digitalen Kodex einen gangbaren Weg darstellt, bestehende oder zu erwartende gesellschaftliche Konflikte zu lösen. Hierfür wurden zunächst Konfliktfelder und der jeweilige Regelungsbedarf identifiziert.

Das Teilprojekt Big Data im Rahmen des Gesamtvorhabens „Braucht Deutschland einen Digitalen Kodex?“ endete nicht mit einem ausformulierten Digitalen Kodex, der abschließend diskutiert wurde. Der folgende Bericht versteht sich vielmehr als Zwischenergebnis in einem Prozess, der weitergeführt werden sollte. Der erste notwendige Schritt lag darin, gemeinsam mit relevanten Akteuren konkrete Konfliktfelder herauszuarbeiten, die durch Big Data entstehen. Da

das Phänomen von Big Data als Regelungsmaterie für einen Kodex insgesamt zu facettenreich ist, wurde die Thematik zunächst auf zwei konkrete Anwendungsfelder für Big-Data-Anwendungen eingegrenzt: Smart Mobility und Smart Health. An ihnen wurden exemplarisch die Chancen und Herausforderungen von Big Data herausgearbeitet, um den Regelungsbedarf zu identifizieren, dem mit einem Digitalen Kodex begegnet werden könnte. Um sich der Thematik anzunähern, war es zunächst erforderlich, den Oberbegriff „Big Data“ näher zu definieren und den Stand der Technik in verschiedenen Anwendungsfeldern zu untersuchen. Anschließend wurden die beiden Unterthemen „Smart Mobility“ und „Smart Health“ auf ihre Potenziale und Risiken hin detailliert beleuchtet.

1.1 Big Data

Wer sich mit dem Regelungsbedarf in Bezug auf Big-Data-Anwendungen befassen will, muss zunächst einmal definieren, was es mit diesem Begriff, der sich in den vergangenen Jahren zu einem Modewort entwickelt hat, eigentlich auf sich hat. Als Ausgangspunkt dient der Begriffsinhalt selbst: Big Data bezeichnet in erster Linie das Vorhandensein großer Datenmengen. Die beinahe alle Lebensbereiche erfassende Digitalisierung der vergangenen drei Jahrzehnte hat zu einer Explosion der vorhandenen Datenmenge geführt. Aufgrund der zunehmenden Komplexität der Anwendungen im Netz gehen derzeitige Prognosen davon aus, dass sich die verfügbare Datenmenge weiter alle zwei Jahre verdoppeln wird. Immer mehr Sensoren und Applikationen sammeln Daten. Beispiele sind Wetterdaten, Verkehrsdaten, Gen-Daten, Textkorpora, Logdateien von Webservern und Nutzerinteraktionen, Bilddaten, die von im öffentlichen Raum installierten Überwachungskameras erzeugt werden, und Daten von RFID-Readern. Hinzu kommen nicht digitale, aber digitalisierbare Daten wie Bücher, analoge Bilder, Fernsehen und Radio. Sie können unstrukturiert sein wie beispielsweise der Text eines Tweets oder eines Blog-Eintrags. Andere sind semistrukturiert

in XML-Dokumenten abgelegt.² Strukturierte Daten sind zum Beispiel Daten aus Tabellen in einer relationalen Datenbank. Viele dieser Daten sind unpräzise und für sich allein wenig aussagekräftig. Wenn sie jedoch massenhaft erzeugt und aggregiert werden, ergibt sich ein enormes Potenzial. Der Begriff „Big Data“ entzieht sich trotz oder gerade wegen seiner ständigen Verwendung in den Medien und im politisch-gesellschaftlichen Diskurs einer festen Definition. Er ist damit offen für die Aufnahme neuer oder erweiternder Bedeutungsnuancen.

„Durch Big Data haben wir nicht nur eine neue Dimension in der Quantität der Daten, sondern auch in der Qualität. Es sind mehr Daten aus ganz verschiedenen Quellen zu vielfältigen Zwecken nutz- und interpretierbar.“

Marit Hansen, Landesbeauftragte für Datenschutz Schleswig-Holstein, Interview

„Das Neue ist, dass wir es im Zeitalter der Digitalisierung nicht mehr nur mit unternehmensinternen Daten zu tun haben, sondern mit Daten aus verschiedenen Quellen. Das erhöht die damit verbundenen Möglichkeiten, aber auch die Komplexität.“

Sven Löffler, Business Development Executive, Big Data & Data-Driven Business T-Systems International GmbH, Konsultation

So ist es wichtig, zu beachten, dass der Begriff nicht nur die „Datenberge“ selbst umschreibt. Er umfasst zugleich die IT-Lösungen und sonstigen computergestützten Methoden, die Institutionen und Unternehmen nutzen, um die verfügbaren Daten auszuwerten, zu analysieren und um Erkenntnisse zur weiteren Verwendung abzuleiten, und die

² XML steht für „Extensible Markup Language“, eine Auszeichnungssprache zur Darstellung hierarchisch strukturierter Daten in Form von Textdateien (Quelle: https://de.wikipedia.org/wiki/Extensible_Markup_Language).

darauf aufsetzenden Abläufe und Geschäftsmodelle. Durch parallelisierte Datenverarbeitung in großen Rechenzentren, z.B. basierend auf den Softwareanwendungen MapReduce und Hadoop, lassen sich die Datenmengen heute größtenteils in Echtzeit verarbeiten. Die algorithmenbasierte Auswertung der Daten führt zu neuen Informationen wie Zusammenhängen und Korrelationen, aus denen sich mitunter sogar Zukunftsprognosen ableiten lassen. Im Vergleich zu früheren Methoden der Datenverarbeitung kehrt sich bei Big Data die Herangehensweise um: Es werden nicht zunächst abstrakt Theorien aufgestellt und die erhobenen Daten anschließend auf Werte und Zusammenhänge hin untersucht, die die Ausgangsthese stützen oder sogar beweisen. Es werden vielmehr die ggf. zunächst für andere Zwecke gesammelten Daten zusammengeführt und dann ausgewertet, um Korrelationen herauszufiltern, die zuvor nicht bekannt waren. Dazu werten die Algorithmen sämtliche verfügbaren Daten, ob strukturiert oder unstrukturiert, aus. Für eine solche Analyse sind die sogenannten Metadaten beispielsweise eines Kommunikationsvorgangs sogar nützlicher als der Inhalt jener Kommunikation selbst. Aufgrund dieser Umkehrung zuvor etablierter Methoden der Wissenschaft wurde im Angesicht des Siegeszugs von Big Data verschiedentlich das „Ende der Theorie“ ausgerufen. Traditionelle Vorgehensweisen seien nicht mehr notwendig dank der Verfügbarkeit und Auswertbarkeit riesiger Datenmengen.

Auch in der universitären oder sonstigen akademischen Forschung ist der Einsatz von Big-Data-Methoden etabliert und allgemein akzeptiert. Die Erhebung von Daten erfolgt hier zumeist nur mit expliziter Einwilligung der Probanden oder Studienteilnehmer. Vorgehensweise und Zielsetzung sind für gewöhnlich transparent, die Ergebnisse werden der Öffentlichkeit zur Verfügung gestellt. Die grundrechtlich abgesicherte Wissenschaftsfreiheit schafft darüber hinaus eine verfassungsrechtliche Grundlage für die Verwendung von Big Data. Anwendungsbereiche, in denen es nicht um die Beobachtung menschlichen Verhaltens geht, sind weitgehend unproblematisch. Konfliktpotenzial birgt Big Data aber dann, wenn es um die massenhafte Erfassung, Speicherung und Auswertung von Daten zu Verhaltensweisen von Menschen geht, vor allem, wenn die Datensammlung nicht für

wissenschaftliche, sondern für wirtschaftliche oder staatliche Zwecke erfolgt.

„Mittlerweile rückt der Datenschutz auch bei Forschern zunehmend ins Zentrum, und es wird überlegt, was berücksichtigt werden muss. Dabei ist vielen Wissenschaftlern nicht bewusst, dass das deutsche Datenschutzrecht durchaus auch Vorschriften enthält, die die Forschung sogar privilegieren. Es ist nicht jegliche Forschung mit personenbezogenen Daten verboten, sondern es gibt besondere Ausnahmenvorschriften, weil Forschung gesellschaftlich wünschenswert ist.“

Dr. Silke Jandt, Vertreterin des Lehrstuhls für Öffentliches Recht, Informationstechnologierecht und Rechtsinformatik an der Universität Passau, Interview

Für sich genommen ist die Erfassung großer Datenmengen keine eigentliche Neuerung. Schon der Datenbestand der bundesrepublikanischen Volkszählung des Jahres 1987 kann als Big Data eingeordnet werden, auch wenn er insgesamt lediglich 80 Megabytes Speicherplatz beanspruchte. Die Bevölkerungs-, Versorgungs- und Verkehrsmitteldaten der Bürger wurden gesammelt, um notwendige infrastrukturelle Maßnahmen einleiten zu können. Damals löste diese, im Vergleich mit heutigen Datenerfassungen vergleichsweise zurückhaltende, Erfassung einen mitunter heftig geführten öffentlichen Diskurs aus. Die Zählung war ursprünglich für 1983 – vier Jahre früher – vorgesehen, musste aber zunächst ausgesetzt werden und wurde anschließend im selben Jahr durch das Bundesverfassungsgericht untersagt. Das Urteil setzte mit der Definition des Rechts auf informationelle Selbstbestimmung neue Maßstäbe für das Datenschutzrecht, die bis in die heutige Debatte um Big Data fortwirken. Dem Gericht zufolge sollen bei der Datenerhebung und -verarbeitung Prinzipien wie Datensicherheit und Zweckbindung stets leitend sein.

„Wir beschäftigen uns oft mit der Frage, wie die deutsche IT-Industrie international und in Bezug auf das Internet wieder einen Fuß auf den Boden bekommen, vielleicht sogar Innovationsführer werden könnte. In diesem Zusammenhang sollte man u.a. darüber nachdenken, ob unsere allgemein eher traditionellen Vorstellungen von Datenschutz, Schutz der Privatsphäre und Urheberrecht im Jahr 2015 noch Bestand haben können oder ob sie im digitalen Zeitalter vielleicht eher hinderlich sind und daher angepasst, vielleicht sogar komplett neu gedacht werden sollten.“

Dr. Georg Rehm, Senior Consultant,
Deutsches Forschungszentrum für
Künstliche Intelligenz GmbH (DFKI),
Konsultation

Angesichts der Entwicklung des Internets und der damit verbundenen umfassenden Digitalisierung sind Datenmengen wie die der Volkszählung vor beinahe 30 Jahren kaum noch der Rede wert. Ein einzelner Nutzer, der pro Tag im Durchschnitt für drei Stunden online ist, lädt alleine dadurch eine jährliche Datenmenge von 65 Gigabytes herunter. Viele der so entstehenden Daten lassen Rückschlüsse auf persönliche Eigenschaften und Vorlieben des Nutzers zu und werden dadurch interessant für unterschiedlichste Akteure wie zum Beispiel Online-Vermarkter oder Marktforschungsunternehmen. Große Teile des globalen Nutzungsverhaltens im Internet werden auf zentralen Servern erfasst und ausgewertet. Diese Daten werden z.B. dafür genutzt, um Werbung gezielt auf den jeweiligen Nutzer abzustimmen und um diese zielgerichteter und damit erfolgversprechender zu gestalten. Neben dem Speichern des reinen Rezeptionsverhaltens entstehen durch die Internet-Nutzung insbesondere umfassende Datenbestände über das Kommunikationsverhalten der Internet-Nutzer. Diese

Daten sind oft zweierlei Natur. Einerseits bestehen sie aus dem eigentlichen Inhalt der Kommunikation. Dieser Inhalt kann interessant sein für jene, die aus den Äußerungen der Nutzer über spezifische Interessen oder Vorlieben Aufschluss über zu schaltende Werbung gewinnen möchten. Andererseits produziert jeder digitale Kommunikationsvorgang sogenannte Metadaten, also beispielsweise, wer mit wem zu welchem Zeitpunkt und für gewöhnlich sogar von welchem Ort aus kommuniziert hat. Diese Daten können für verschiedene Akteure äußerst aufschlussreich und daher von großem Wert sein.

„1,4 Milliarden Menschen werden über den Facebook-Newsfeed informiert. Es reicht, wenn man hier nur ein Prozent beeinflusst. Das kann riesige Auswirkungen haben. Hier ist die eigentliche Gefahr von Big Data.“

Michael Seemann, Kulturwissenschaftler,
Autor und Blogger, Konsultation

Versuche haben gezeigt, dass sich aus der Analyse allein der Kommunikationsmetadaten von Arbeitnehmern die Kündigungswahrscheinlichkeit einzelner Personen ableiten lässt. Auf diese Weise kann der Arbeitgeber abschätzen, ob ein Arbeitnehmer dem Betrieb erhalten bleibt oder eigentlich bereits auf dem Absprung ist – und das möglicherweise schon, bevor es dem Arbeitnehmer selbst vollends bewusst ist.³

Neben der Wirtschaft entwickeln staatliche Stellen ein zunehmendes Interesse an der Auswertung von Metadaten. Wie groß die Begierde mitunter ist, haben nicht zuletzt die Enthüllungen des Whistleblowers Edward Snowden im Jahre 2013 gezeigt. Berichten und realistischen Einschätzungen zufolge ist der amerikanische Geheimdienst NSA auf dem Weg, jedenfalls in bestimmten, kleineren Ländern sämtliche digital vermittelten menschlichen Kommunikations- und Rezeptionsvorgänge zu erfassen und zu speichern.⁴

³ Siehe Frank Rieger, „Der Mensch wird zum Datensatz“, faz.net, 16. Januar 2010, <http://www.faz.net/aktuell/feuilleton/ein-echtzeitexperiment-der-mensch-wird-zum-datensatz-1591336.html>.

⁴ Siehe z.B. B. Gellman und A. Soltani, „NSA surveillance program reaches 'into the past' to retrieve, replay phone calls“, in Washington Post, 18. März 2014, https://www.washingtonpost.com/world/national-security/nsa-surveillance-program-reaches-into-the-past-to-retrieve-replay-phone-calls/2014/03/18/226d2646-ade9-11e3-a49e-76adc9210f19_story.html.

„Bei allen Daten, die wir erhalten, zum Beispiel aus Volkszählungen, achten wir immer darauf, dass man keine Rückschlüsse auf das Individuum ziehen kann.“

Sabrina Juran, Technical Specialist,
Data and Research, United Nations
Populations Fund (UNFPA), Interview

In jüngerer Zeit hat sich zu den Kommunikationsdaten im Internet eine neue Kategorie hinzugesellt, deren Potenzial sich erst allmählich zu entfalten beginnt: jene Daten, die durch Geräte erzeugt werden, die sich im weitesten Sinne dem sogenannten Internet der Dinge zuordnen lassen. Darunter fallen vernetzte Geräte wie zum Beispiel moderne Hausmesstechnik und -steuerung, Fahrzeuge und sogenannte Wearables. Hier sind zurzeit erhebliche Innovationschübe in Richtung einer umfassenden Digitalisierung und Vernetzung zu beobachten, die die Möglichkeiten des Sammelns von Nutzerdaten noch einmal drastisch ausweiten.

So werden beispielsweise über die in den Geräten privater Heizungstechnik enthaltenen Sensoren detaillierte Daten über die Bewohner eines Hauses ermittelt: wann diese zu Hause sind, in welchem Zimmer sie sich aufhalten, welches ihre bevorzugte Raumtemperatur ist und allerhand mehr. Die Vorteile für den Nutzer liegen auf der Hand. Die darauf basierenden Innovationen versprechen unter anderem erhebliche Einsparungen für den Energieverbrauch.

Das Thermostat der amerikanischen Firma Nest ist das bislang wohl bekannteste Beispiel für „smarte“ Haustechnik. Es sammelt neben den Angaben über den Ort der Installation – sofern der Nutzer diesen wie empfohlen angibt – Sensor- und Nutzungsdaten des Geräts und übermittelt sie an die Server des Unternehmens. Der Datenschutzerklärung zufolge misst das Thermostat Raumtemperatur, Luftfeuchtigkeit und Lichteinwirkung. Außerdem meldet ein eingebauter Bewegungsmelder, ob sich etwas im Raum bewegt. Dabei lernt das Thermostat aus dem Verhalten der Nutzer. Das Unternehmen verspricht, dass dadurch der Verbrauch von Heizung und Klimatechnik

um bis zu 26 Prozent reduziert werden könne. Weitgehend unbekannt ist, wie Nest die entstandenen und übermittelten Daten verarbeitet. Wie sehr allerdings der Wert eines Unternehmens steigt, wenn durch den vermeintlich primären Anwendungszweck eine große Menge an Nutzerdaten generiert, gespeichert und zur Weitergabe zur Verfügung gestellt wird, macht die Tatsache deutlich, dass Nest Anfang 2014 für 3,2 Milliarden Dollar von Google übernommen wurde.

Weitere Anwendungsfälle für das Internet der Dinge sind neben Haushaltsgeräten insbesondere „intelligente“ vernetzte Fahrzeuge sowie Wearables wie Smart Watches oder Fitnessarmbänder, auf die sich der Bericht im Folgenden fokussieren wird. Ein im Juli 2015 erschienenenes White Paper des „Forums Privatheit“ spricht in diesem Zusammenhang vom „versteckten Internet“. Dies spielt auf den Umstand an, dass es sich um Geräte handelt, in welche die vernetzten Computer so integriert sind, dass sie nicht mehr als Computer wahrgenommen werden.⁵ Für das Thema Big Data ist dieser Umstand zentral. Denn der Nutzer nimmt häufig nicht mehr wahr, dass es sich bei den Geräten um Daten sammelnde Computer handelt und dass mit ihnen unter Umständen sensible Daten erzeugt werden, die dann an den Anbieter des Geräts oder sogar an Dritte übermittelt werden. Diesem Produktbereich kommt daher eine besondere Relevanz zu, da hier völlig neue, bisher unerschlossene Datenquellen zugänglich gemacht werden, die zugleich einen sehr persönlichen Charakter aufweisen.

„Gerade in Zeiten von Big Data werden die Anonymisierungsmöglichkeiten immer geringer. Trotzdem sollte es das Ziel sein, den Personenbezug so weit wie möglich auszudünnen, um möglichst viele Big-Data-Anwendungen auch gesellschaftlich akzeptabel zu gestalten.“

Dr. Alexander Dix, LL.M., Berliner
Beauftragter für Datenschutz und
Informationsfreiheit,
1. Expertenworkshop, 26.02.2015

⁵ Forum Privatheit (Hg.), „Das versteckte Internet. Zu Hause – Im Auto – Am Körper“, White Paper, https://www.forum-privatheit.de/forum-privatheit-de/aktuelles/aktuelles_dokumente/White-Paper-2-Final_17.07.15-Druckversion.pdf.

INTERVIEW MIT ANDREA VOSSHOFF

Big Data – Chancen und Risiken

? Die gesellschaftlichen Chancen und Risiken von Big-Data-Anwendungen werden zunehmend diskutiert. Wie nehmen Sie in Ihrer Arbeit diese Debatte wahr? Können Sie ein verstärktes Problembewusstsein wie auch einen höheren Bedarf an der Klärung von Fragen in diesem Bereich in ihrer täglichen Arbeit ausmachen?

Andrea Vosshoff: Einerseits merke ich in meiner täglichen Arbeit schon, dass das Bewusstsein der Bürger für die Risiken, die sich aus Big-Data-Anwendungen ergeben, gewachsen ist. Gleichzeitig ist zu beobachten, dass immer mehr Menschen bereit sind, ihre Daten für einen vermeintlich finanziellen Vorteil oder sonstige Annehmlichkeiten preiszugeben, ohne dass sie wissen oder sich konkret informieren können, was mit ihren Daten geschieht.

? Was ist vor dem Hintergrund dieser Entwicklungen aus Ihrer Sicht das Neue im Bereich Tracking und Big Data, wenn es um den digitalen Raum geht? Was hat sich im Vergleich zu analogen Zeiten geändert?

AV: Im digitalen Zeitalter werden immer mehr Daten generiert. So ist zu lesen, dass das weltweite digitale Datenvolumen im Jahr

Andrea Voßhoff

Geboren 1958 in Haren/Ems, ist Andrea Voßhoff seit Januar 2014 die Bundesbeauftragte für den Datenschutz und die Informationsfreiheit.

Sie studierte Rechtswissenschaften an der Universität Münster und absolvierte ein Auslandsstudium an der Universität von Lausanne. Nach ihrem zweiten juristischen Staatsexamen arbeitete sie zu-

nächst als Rechtsanwältin in Haren/Ems. Später war sie in einem Notarbüro in Rathenow tätig. Von 1998 bis 2013 war Andrea Voßhoff Mitglied des Deutschen Bundestages und dort ordentliches Mitglied im Rechtsausschuss. Von 2010 bis 2013 war sie außerdem rechtspolitische Sprecherin der CDU/CSU-Bundestagsfraktion.

Foto: Bundesregierung/Kugler



Problematisch ist auch, dass etwa Betreiber nicht immer ausreichend über Sinn und Zweck der Datenerhebung Auskunft geben.

Andrea Voßhoff

2012 auf 7,8 Zettabytes geschätzt wurde, eine Zahl mit 21 Nullen. Wegen immer leistungsfähigerer Speicher- und Rechnertechnologien können diese umfassend gesammelt und nahezu in Echtzeit ausgewertet werden. Es besteht zudem die Gefahr, dass vermeintlich anonyme Daten durch die Zusammenführung verschiedenster Informationen aus unterschiedlichen Quellen wieder personenbeziehbar werden können. Liegen bereits personenbezogene Daten vor und werden diese miteinander verknüpft, so ist die einfache Erstellung von sehr persönlichen Profilen möglich. Problematisch ist auch, dass etwa Betreiber nicht immer ausreichend über Sinn und Zweck der Datenerhebung Auskunft geben und Nutzer nicht über die AGB informiert sind und somit nicht einmal wissen, wer was über sie gespeichert hat und was mit den Daten passiert.

? In welchen Situationen sind Sie privat wie auch beruflich beunruhigt, wenn Sie von Datensammlungen durch Privatunternehmen hören?

AV: Ich bin insbesondere dort beunruhigt, wo zentrale Datenschutzgrundsätze wie Zweckbindung und Datensparsamkeit nicht beachtet werden und die Verwendung der Daten für den Nutzer nicht wirklich transparent ist. Wenn dann noch das Unternehmen seine Marktmacht ausspielt und dem Bürger keine umfassende und informierte Einwilligung in die Nutzung seiner Daten ermöglicht, gibt dies Anlass zur Sorge.

? Wie schätzen Sie die Risiken ein, die sich aus der Erstellung von Profilen über Einzelpersonen mithilfe von Sammlungs-, Tracking- und Analysetechniken im Bereich von Big Data ergeben können?

AV: Big-Data-Technologien sind auch eine Grundlage für (verbesserte) Scoring-Methoden, d.h. die Vorhersage menschlichen Verhaltens auf der Basis mathematisch-statistischer Auswertungen großer Datenbestände. Hier besteht das Risiko der Kategorisierung jedes Einzelnen, der man sich nur schwer entziehen kann,

wenn die Vorhersage des Verhaltens nicht der Realität entspricht. Die Nutzung von Big-Data-Technologien und die darauf basierende Profilbildung erhöhen den Anpassungsdruck zu konformem und unauffälligem Verhalten.

? Wie ist es bei privaten Akteuren – wie sehen Sie den Unterschied zwischen der Datensammlung der Schufa und der Datensammlung von Facebook?

AV: Sowohl bei der Schufa als auch bei Facebook handelt es sich um nicht öffentliche Stellen im Sinne des Bundesdatenschutzgesetzes, die in großem Umfang personenbezogene Daten erheben und verarbeiten. Während aber für die Weitergabe von Daten an Auskunftsteile und deren dortige Verarbeitung spezifische gesetzliche Regelungen zu beachten sind, fehlen diese oftmals für Soziale Netzwerke. Zudem ist es höchst zweifelhaft, ob deutsches Datenschutzrecht auf US-amerikanische Unternehmen überhaupt anwendbar ist. Wer solche Dienste nutzen will, hat oft keine wirklich selbst-



BIG DATA – CHANCEN UND RISIKEN

“ In einer globalisierten digitalen Welt, in der Daten über das Internet in Echtzeit überall verfügbar sind, kann nationaler Schutz allein nicht ausreichend sein.

Andrea Voßhoff



bestimmte Entscheidungsbefugnis hinsichtlich seiner Nutzungsbedingungen, sondern kann diesen nur pauschal zustimmen.

? Welche Daten sollten der Wirtschaft vorenthalten werden? Gibt es Daten, die Sie für zu sensibel halten, um sie für die Nutzung durch Big-Data-Anwendungen zu öffnen?

AV: Das ist kaum so pauschal zu beantworten. Grundsätzlich gibt es im Big-Data Zeitalter keine „harmlosen“ Daten mehr. Die besonderen Risiken liegen ja gerade in der Kombination vieler scheinbar „nichtssagender“ Daten zu einem umfassenden Profil.

? Ist Ihrer Ansicht nach eine (national)staatliche Regulierung ausreichend, oder bedarf es einer besseren europäischen bzw. internationalen Kooperation – oder einer Mischung daraus?

AV: Ich habe seit Beginn meiner Amtszeit immer wieder betont: In einer globalisierten digitalen Welt, in der Daten über das Internet in Echtzeit überall verfügbar sind, kann nationaler Schutz allein nicht ausreichend sein. Wo der Datenaustausch grenzenlos ist, muss auch der Schutz international verankert sein.

? Gehen wir mit der Datenschutzgrundverordnung

auf europäischer Ebene den richtigen Weg? Sind Sie zufrieden mit den aktuellen Regelungsvorschlägen darin, oder sollte aus Ihrer Sicht an der einen oder anderen Stelle durch den Verhandlungspartner Deutschland versucht werden, noch nachzubessern?

AV: Die Datenschutzgrundverordnung ist der richtige Weg, wenn es darum geht, auch beim Datenschutz ein einheitliches Niveau im Vereinten Europa zu schaffen. Im Ergebnis darf die Reform des Europäischen Datenschutzrechts nicht dazu führen, hinter dem geltenden Datenschutzniveau der EG-Datenschutzrichtlinie aus

dem Jahre 1995 zurückzubleiben. Die sich aus Artikel 8 der Grundrechtecharta und Art. 16 Abs. 1 AEUV ergebenden Grundprinzipien des Datenschutzes dürfen nicht zur Disposition stehen. Ich appelliere daher an alle Beteiligten, sich auf die Werte zu besinnen, die die Grundsätze wie Zweckbestimmung, Einwilligung und Datensparsamkeit im Kontext der informationellen Selbstbestimmung darstellen. Hierbei geht es nicht um beliebige Details, sondern um Kernbereiche des Datenschutzes, die in Zeiten der fortschreitenden Globalisierung und umfassenden Vernetzung der Datenverarbeitung von wachsender Bedeutung sind. In dem Entwurf der Datenschutzgrundverordnung fehlen spezifische Anforderungen an riskante Datenverarbeitungen wie zum Beispiel Profilbildungen und -nutzungen oder Videoüberwachung. Auch sollen Daten für Werbezwecke weiterhin ohne Einwilligung der Betroffenen verarbeitet werden können.

Gerade in Zeiten von Big Data und globaler Datenverarbeitung sind die Autonomie des Einzelnen, Transparenz und Rechtmäßigkeit der Datenverarbeitung, die Zweckbindung und die Verantwortlichkeit des Datenverarbeiters ebenso wichtige Elemente der Grundrechtsgewährleistung wie

eine starke Datenschutzaufsicht und wirksame Sanktionen. Als positiv betrachte ich die sowohl vom Europäischen Parlament als auch vom Rat gewollte Stärkung der Zusammenarbeit der europäischen Datenschutzbehörden innerhalb des sogenannten „One-Stop-Shops“ und des künftigen Europäischen Datenschutzausschusses. Der effiziente Vollzug des Datenschutzrechts darf jedoch nicht durch die Untätigkeit der federführenden Datenschutzbehörde unterlaufen werden. Es sollte daher eine Regelung geschaffen werden, wonach die mitgliedstaatlichen Aufsichtsbehörden bei Betroffenheit ihrer Bürger von der federführenden Behörde ein aufsichtsbehördliches Einschreiten verlangen können.

? Ist eine eventuelle Regulierung der wirtschaftlichen Nutzung von Big Data ein Thema in der Artikel 29-Datenschutzgruppe?

AV: Die Art. 29 Working Party hat sich anlässlich des 97. Plenums am 16./17.09.2014 mit Big Data befasst und ein Statement zu den Auswirkungen von Big Data auf den Betroffenenenschutz verabschiedet. Auch die Internationale Konferenz der Beauftragten für den

Datenschutz und die Privatsphäre hat sich mit den datenschutzrechtlichen Herausforderungen von Big Data befasst.

? Big Data birgt nicht nur Risiken, sondern bietet sowohl der Gesellschaft wie auch dem Individuum viele Vorteile. Können Sie hierzu Beispiele nennen, die Sie besonders eindrucksvoll finden?

AV: Big Data bietet Chancen für gezielte Ressourcenverteilung der Energiewirtschaft, in der medizinischen Forschung, der Früherkennung von Epidemien oder Katastrophen oder etwa in der Marktforschung. Daneben kann die Big-Data-Technologie helfen, Trends frühzeitig zu erkennen und entsprechend gezielt zu agieren. Dies ist zu begrüßen, sofern es in Einklang mit dem Datenschutz geschieht. Es besteht jedoch stets das Risiko der Kategorisierung des Einzelnen, der er sich selbst dann nur schwer entziehen kann, wenn die durch Big-Data-Anwendungen getroffene Vorhersage seines Verhaltens nicht der Realität entspricht. Big-Data-Projekte müssen sich deshalb im vorgegebenen rechtlichen Rahmen bewegen. Es bedarf darüber hinaus eines fortwährenden gesellschaftspoli-



BIG DATA – CHANCEN UND RISIKEN



tischen Diskurses über das Verhältnis zwischen technologischem Fortschritt und notwendiger rechtlicher Regulierung.

? Wie kann man aus Ihrer Sicht die Vorteile von Big Data noch besser fördern? Brauchen wir nicht auch ein „Big-Data-Chancengesetz“?

AV: Meine Aufgabe ist es, mich für die Chancen des Datenschutzes starkzumachen.

? Was wünschen Sie sich ganz persönlich von den Beteiligten in der Debatte um die Ausgestaltung und Nutzung von Big Data? Brauchen wir mehr Geschwindigkeit, mehr Ruhe oder auch ganz andere Blickwinkel?

AV: Ich wünsche mir im Zusammenhang mit der Nutzung von Big Data eine stärkere wertebegleitete Debatte hinsichtlich der gesellschaftlichen Folgen. Wie immer, wenn es um neue Technologien geht, sollte daher auch hier bedacht werden: Nicht alles, was machbar ist, muss auch sinnvoll sein. Anwendungen können nicht allein deswegen zugelassen werden, weil sie technisch möglich sind.

? Haben Sie einen Vorschlag, wie Privatsphäre und Datenschutz mit einer effektiven Nutzung von Big Data vereinbar sind?

AV: Das wäre dort der Fall, wo Big-Data-Anwendungen ausschließlich mit zuverlässig anonymisierten Daten arbeiten und auch durch das Zusammenführen massenhafter und heterogener Daten keine Repersonalisierung möglich ist. Es fehlen zudem wirksame Instrumente, z.B. Gesetze und Sanktionen, um nachhaltig den Datenschutz bei Big Data einzufordern.

? Wie stehen Sie als Bundesbeauftragte für den Datenschutz und die Informationsfreiheit zu der Idee selbst fahrender Autos?

AV: Die Idee ist durchaus faszinierend, auch wenn ich mein Auto gern selbst lenke.

? Wer soll wissen dürfen, was Sie wo und wann mit Ihrem Auto machen?

AV: Ich und – wenn ich es für erforderlich halte – die Werkstatt meines Vertrauens.

? Ist es in Ordnung, dass Versicherungen aus Fahrdaten der Versicherungsnehmer bestimmte Schlussfolgerungen ziehen könnten, die sich auf deren Vertrag auswirken?

AV: Ein Versicherungstarif, der auf Daten zum Fahrverhalten der Versicherungsteilnehmer basiert, ist aus datenschutzrechtlicher Sicht ablehnend zu bewerten.

Auch wenn die Freiwilligkeit einer solchen Tarifstruktur betont wird, eine entsprechende Prämienengestaltung durch immer mehr Versicherer würde zu einem „ökonomischen Zwang“ zur Wahl solcher Tarife führen.

Wer einen solchen Tarif wählt, sollte sich bewusst sein, dass eine Vielzahl von personenbezogenen Daten erhoben, gespeichert und ausgewertet wird. Und er sollte wissen, von wem und wofür.

? Nutzen Sie eine App für die Aufzeichnung und Überwachung von Gesundheitsdaten? Wie sehen Sie solche Apps in Ihrer Funktion als Bundesbeauftragte für den Datenschutz und die Informationsfreiheit?

Ich warne daher ohne entsprechende Transparenz vor den datenschutzrechtlichen Risiken und Nebenwirkungen solcher Apps.

Andrea Voßhoff

AV: Ich nutze keine solche App und lehne die Nutzung solcher Anwendungen aus datenschutzrechtlichen Gründen für mich ab.


? Sehen Sie es grundsätzlich als Vorteil oder Nachteil, dass Gesundheitsdaten wie etwa die tägliche körperliche Aktivität oder die Herzaktivität über die Nutzer bekannt sind?

AV: Soweit derartige Apps dazu genutzt werden, um etwa akut gefährdeten Patienten zu helfen, ist sicherlich nichts gegen derartige Gesundheits-Apps einzuwenden. Dies gilt jedoch nur für einen äußerst geringen Teil dieser Apps. So ist gegen die Übermittlung etwa von Blutzuckerwerten bei Diabetikern oder Blutdruckwerten bei Herzpatienten unmittelbar an den Arzt nichts einzuwenden. Möglicherweise werden hier auch Leben gerettet. Gerade bei der telemedi-

zinischen Überwachung von Herzpatienten erwarten Studien eine Erhöhung der Überlebensraten von bis zu 15 Prozent. Es handelt sich hierbei aber in der Regel um Apps, die nach § 3 Medizinproduktegesetz ein Konformitätsverfahren durchlaufen haben. Das ist aber die große, große Ausnahme.

Bei den allermeisten sogenannten Health-Apps handelt es sich eher um Lifestyle-Apps mit einem medizinischen Bezug, und da stellt sich die Frage: Wem werden diese sensiblen personenbezogenen Daten überhaupt bekannt? Dem App-Entwickler? Dem Vertreiber der App? Dem Sportverband, dem Arzt oder der privaten oder gesetzlichen Krankenkasse? Und zu welchem Zweck werden diese Daten erhoben und verwendet? Ich warne daher ohne entsprechende Transparenz vor den datenschutzrechtlichen Risiken und Nebenwirkungen solcher Apps.

? Finden Sie es in Ordnung, wenn Versicherungen diese Daten auswerten können und die Ergebnisse mit ihren Vertragskonditionen verknüpfen?

AV: Nein, auch wenn es Menschen, die jung und gesund sind (oder sich gesund fühlen), vorteilhaft erscheint, durch Bereitstellung ihrer Daten einen günstigeren Tarif oder andere Vorteile erhalten zu können: Durch die unbedachte Übermittlung von Gesundheitsdaten, die nach unbekanntem Algorithmen ausgewertet werden, entsteht für den Versicherten ein unkalkulierbares datenschutzrechtliches Risiko. Zudem wächst der Druck, in die Übermittlung sensibler Gesundheitsdaten einzuwilligen, auch für den, der dies gar nicht möchte, aber befürchten muss, ansonsten keine private Kranken- oder Lebensversicherung mehr abschließen zu können. 

1.2 Tracking

Big Data ist auf die Gewinnung möglichst vieler Einzeldaten angewiesen – und das Internet der Dinge eröffnet vorher ungekannte Möglichkeiten der Erzeugung eben solcher Daten. Insbesondere, wenn es sich dabei um bewegliche Geräte handelt, wird die Methode des Sammelns nutzerbezogener Daten als Tracking bezeichnet. Die Methode, Big Data auf diese Weise zu generieren, steht im Folgenden im Fokus des vorliegenden Berichts.

Das englische Wort „Tracking“ lässt sich mit Folgen, Verfolgen oder Nachverfolgen übersetzen. Es beschreibt all jene Handlungsschritte, die notwendig sind, um die Bewegung einer Sache oder einer Person in Echtzeit nachzuvollziehen. Tracking ermöglicht es, den Aufenthaltsort⁶ eines Objekts zu jedem Zeitpunkt fortlaufend zu bestimmen. Dies grenzt Tracking vom sogenannten Tracing ab, worunter das zeitlich versetzte, also nachträgliche Verfolgen eines Bewegungsablaufes verstanden wird. Allerdings werden diese Begriffe nicht immer trennscharf verwendet. Vielmehr kommen sie oft als feststehendes Begriffspaar vor und bezeichnen dann gemeinsam und ohne weitere Unterscheidung beide Formen des Verfolgens, das gleichzeitige wie auch das nachträgliche.

„Es wird so viel über eine Blackbox im Auto diskutiert. Das ist widersinnig, wenn man bedenkt, dass bei fast jedem eigentlich eine Blackbox in Form eines Handys in der Jackentasche mitläuft.“

Ivo Körner, Geschäftsführer Vertrieb
Branchenkunden IBM Deutschland GmbH,
öffentliche Veranstaltung Stuttgart,
25.06.2015

Traditionell wird der Begriff Tracking im Zusammenhang mit dem Nachverfolgen von Waren verwendet. Er bezieht sich in diesem Sinne in erster Linie auf Bewegungs- und Ortungsdaten. In der Logistik

werden schon lange Tracking-Systeme eingesetzt, um den Aufenthaltsort versendeter Sachen jederzeit bestimmen zu können. Dazu dienen einerseits Barcodes oder an den Objekten angebrachte RFID-Chips⁷, die an bestimmten, vorher festgelegten Punkten gescannt werden, so beispielsweise am Ausgang des Warenlagers. Während des Transportes selbst können die Objekte durch den Einsatz von in Transportmitteln installierten GPS-Chips⁸, die ihren Aufenthaltsort per Satellit bestimmen, verfolgt werden.

“Die Nutzung von Big-Data-Technologien und die darauf basierende Profilbildung erhöhen den Anpassungsdruck zu konformem und unauffälligem Verhalten.“

Andrea Voßhoff, Bundesbeauftragte für den Datenschutz und die Informationsfreiheit, Interview

Heute lässt sich auch der Aufenthaltsort von Personen fast lückenlos in Echtzeit bestimmen. Noch vor wenigen Jahren musste man zum Erstellen von Bewegungsprofilen – zum Beispiel im Rahmen von Ermittlungen in Strafverfahren – noch mit wenigen Ortsdaten auskommen und war darauf angewiesen, dass die betreffende Person ihre Kreditkarte benutzte oder von Zeugen auf Aufnahmen von Überwachungskameras erkannt wurde. Heute stellt das Tracking von Personen durch die praktisch vollständige Verbreitung von Mobiltelefonen für Strafverfolgungsbehörden oder Geheimdienste, aber auch für private Unternehmen kein größeres Problem mehr dar. Neben der Ortung per GPS-Chip, der in fast allen neueren Smartphones eingebaut ist, damit der Besitzer die standortbezogenen Dienste (sogenannte Location-based Services) nutzen kann, ermöglicht es insbesondere die Einwahl des Telefons in die jeweils nächstgelegene Funkzelle oder das Einloggen in ein bereitgestelltes WLAN, umfassende Bewegungsprofile zu erstellen. Die Detailfülle solcher Profile wurde vor einigen Jahren vom Politiker Malte Spitz aufgezeigt, nachdem er Zugang

⁶ Eine Erweiterung auf die Verfolgung anderer Verhaltensparameter erfolgt später im Text.

⁷ RFID steht für „Radio Frequency Identification“ und bezeichnet eine Technologie für Sender-Empfänger-Systeme zum automatischen und berührungslosen Identifizieren und Lokalisieren von Objekten und Lebewesen mit Radiowellen (Quelle: <https://de.wikipedia.org/wiki/RFID>).

⁸ GPS steht für „Global Positioning System“, ein globales Navigationssatellitensystem zur Positionsbestimmung (Quelle: https://de.wikipedia.org/wiki/Global_Positioning_System).

zu den über ihn gespeicherten Vorratsdaten bei der Deutschen Telekom eingeklagt hatte.⁹

Ein konkretes Beispiel für das Tracking von Personen, analog zur Nachverfolgung von Warenbewegungen in Echtzeit, ist das Tracking von Kindern mittels Armbändern, die mit RFID-Chips ausgestattet sind. Diese wurden schon 2004 im Freizeitpark Legoland im dänischen Billund eingeführt, damit Eltern ihre Kinder innerhalb des Parks orten und wiederfinden können.

Das Bestimmen der Ortsdaten von Sachen oder Personen in Echtzeit kann als Tracking im engeren Sinne umschrieben werden. Jedoch hat es sich durchgesetzt – insbesondere, seit moderne Mobilfunktechnologien weitreichend verfügbar sind –, ganz allgemein die nachverfolgende Erhebung unterschiedlichster Einzeldaten und Parameter als Tracking im weiteren Sinne zu bezeichnen. Dazu gehört unter anderem das sogenannte Activity Tracking, also die Aufzeichnung einzelner Fitness- oder Gesundheitsdaten eines Nutzers.¹⁰ Aber auch das Dokumentieren von Fahrtparametern wie Geschwindigkeit, Bremsverhalten oder Beschleunigung mittels eines in einem Fahrzeug installierten Gerätes, das die Daten per Funktechnologie auf den Server eines Versicherungsunternehmens übermittelt, wird inzwischen unter den Begriff Tracking gefasst, obwohl es hier zu meist nicht um eine Verfolgung des Verlaufs der Einzeldaten in Echtzeit geht, sondern um das nachträgliche Nachvollziehen des vorangegangenen Prozesses.

„Ein Problem mit vernetzten Gegenständen ist, dass die Menschen sich oft gar nicht bewusst sind, dass sie getrackt werden. Die Nutzer sollten auf verständliche Weise darüber informiert werden, was geschieht.“

Prof. Dr. Tobias O. Keber, Professor für Medienrecht und Medienpolitik, Institut für Digitale Ethik (IDE) der Hochschule der Medien Stuttgart, Interview

Trotz solcher begrifflicher Unschärfen kann man unter Tracking ganz allgemein die kontinuierliche Aufzeichnung bestimmter Einzelinformationen in ihrem zeitlichen Verlauf verstehen, die sich in der Ge-

samtbetrachtung zu einem erkenntniserweiternden Datensatz zusammensetzen. Das kann die Wegstrecke sein, die eine Person in einer bestimmten Zeitspanne zurückgelegt hat, die Entwicklung der Herzfrequenz, während die Person einen Halbmarathon absolviert, oder die Häufigkeit scharfer Bremsungen während einer Autofahrt von Hamburg nach Berlin.

Die so erzeugten individuellen Datensätze umschreiben zunächst lediglich einen abgegrenzten Lebenssachverhalt, aus dem der Nutzer des Geräts relevante Erkenntnisse für sich ziehen kann. Isoliert betrachtet handelt es sich bei den entstandenen Werten um „Small Data“, also um Datensätze, die sehr spezifische, für den Einzelnen unmittelbar verständliche und für bestimmte oder bestimmbar Einsatzgebiete nützliche Informationen bereithalten. Werden viele dieser Datensätze zusammengeführt, kann aus ihnen Big Data entstehen. Die Auswertung und Verarbeitung von Big Data ermöglicht neue Erkenntnisse, die sich aus den Einzeldaten allein nicht hätten gewinnen lassen.

Woraus diese einzelnen Datensätze gewonnen werden, ist für die Einordnung nebensächlich. Allerdings ist mit der Entwicklung von entsprechenden Technologien und Geräten auf unterschiedlichen Gebieten Tracking zu einer der vorrangigen Methoden der Datengewinnung geworden, die in den kommenden Jahren mit aller Wahrscheinlichkeit gesellschaftlich noch relevanter werden wird. Schon aus diesem Grund bietet es sich an, im Hinblick auf die Frage nach der Notwendigkeit eines Digitalen Kodex für Big Data diesen Aspekt speziell zu beleuchten.

„Es wird immer angenommen, wir sind gerade in den Prozess der Digitalisierung eingestiegen. Das stimmt aber nicht: Wir sind schon lange in diesem Prozess, und in manchen Bereichen sind wir auch kurz vor dem Quantensprung.“

Minister für Verkehr und Infrastruktur Baden-Württemberg, **Winfried Hermann**, öffentliche Veranstaltung Stuttgart, 25.06.2015

⁹ Die Vorratsdatenspeicherung betrifft allerdings eher den Aspekt des Tracing, also die nachträgliche Nachverfolgung der Bewegung von Personen.

¹⁰ Vgl. https://de.wikipedia.org/wiki/Activity_Tracker.

1.3 Entwicklung und Stand der Technik

Datenanalysen kommen in vielen Anwendungsfeldern zum Einsatz. Ebenso vielfältig sind die Verfahren, die zu diesem Zweck verwendet werden, beispielsweise „Query and Reporting“ (Suchen und Berichten), „Data-Mining“¹¹, Datenvisualisierung, Vorhersagemodelle und Prognosen, Optimierung (vor allem von Prozessen), Simulation, Integration verschiedener Datenformate (etwa bei Sprach- und Bilderkennung), Geodaten- und raumbezogene Analysen.

Wie bereits erwähnt, bezieht sich der Begriff „Big Data“ nicht nur auf die großen Datenmengen, die durch vernetzte Technologie erzeugt werden. Darüber hinaus ist die Verarbeitung dieser Daten selbst einbezogen. Dieser Vorgang wird gemeinhin mit den „drei V“ umschrieben: Volume (Volumen), Velocity (Geschwindigkeit) und Variety (Vielfalt). Als grundlegende Methode der Auswertung kommt die Wahrscheinlichkeitsrechnung hinzu, insbesondere als Verfahren der Wissensgenerierung von Computern (sogenanntes Machine Learning – maschinelles Lernen).

Volume

Dank moderner technischer Architekturen ist es heute möglich, auch sehr große Datenmengen in überschaubarer Zeit zu verarbeiten. Im Unterschied zu den traditionellen Datenbankmodellen erlauben es die neuen Verfahren, die Verarbeitung von Daten auf parallel arbeitenden Computern zu koordinieren. Erst dadurch werden die den Big-Data-Anwendungen zugrunde liegenden Datenmassen handhabbar. Ein Beispiel ist der Suchindex von Google, der gegenwärtig eine geschätzte Größe von 100 Petabytes (das sind 100 Millionen Gigabytes) aufweist – und er wächst weiter, da immer neue Webseiten indiziert und gespeichert werden. Um die mehreren Milliarden Suchanfragen, die jeden Tag eingegeben werden, zu beantworten, ist die gleichzeitige Bearbeitung auf mehreren Computern notwendig. Diese Parallelisierung ist technisch anspruchsvoll und hat sich daher erst in den letzten Jahren am Markt durchgesetzt. In-

zwischen ist dieses Verfahren so weit etabliert, dass es im Rahmen von Cloud-basierten Anwendungen im Netz für jeden nutzbar ist.

Velocity

Auch für den Verarbeitungsaspekt der Geschwindigkeit mag Google als Beispiel dienen. Die Zeit, die für die Verarbeitung großer Datenbestände notwendig ist, stellt nach wie vor einen limitierenden Faktor dar. Denn ein Ergebnis ist oft nur dann hilfreich, wenn es auch sehr schnell verfügbar ist. Hierauf basiert letztlich das Geschäftsmodell des weltgrößten Suchmaschinen-Anbieters. Soll neben der Beantwortung einer Suchanfrage gleichzeitig die passende Werbung eingeblendet werden und auf das aktuelle Surfverhalten abgestimmt sein, dann stellt dies sehr hohe Anforderungen an die Verarbeitungsgeschwindigkeit. Wenn die Auswahl der richtigen Werbung nicht rechtzeitig für das entsprechende Suchergebnis kommt, wird sie vom Nutzer kaum oder gar nicht wahrgenommen, und der Vorgang ist für den Werbekunden wertlos. Verzögert sich aber andererseits die Beantwortung der Suchanfrage auch nur um Bruchteile einer Sekunde, besteht die Gefahr, dass die Suchmaschine Marktanteile bei den Nutzern verliert. Mit der Verbreitung von sogenannten Real-Time-Bidding-Systemen¹², die es ermöglichen, automatisiert im Augenblick der getätigten Suchanfrage für die Platzierung einer Anzeige ein Gebot abzugeben, gibt es einen erheblichen Bedarf an Modellen, die den Wert eines solchen schnell auf einer großen Datenbasis errechnen können.

Variety

Die Struktur der vorhandenen Daten stellt nach wie vor eine Herausforderung für die Verarbeitung dar. Einzelne Datenkomponenten in Datenbanken sind stets durch einen Feldtyp näher beschrieben, so zum Beispiel Vorname, Nachname oder das Geburtsdatum. Den größten Teil aller vorhandenen Datenbestände wird man demgegenüber als unstrukturiert bezeich-

¹¹ Unter Data-Mining versteht man die systematische Anwendung statistischer Methoden auf große Datenbestände mit dem Ziel, neue Querverbindungen und Trends zu erkennen (Quelle: <https://de.wikipedia.org/wiki/Data-Mining>).

¹² Real-Time-Bidding bedeutet Echtzeit-Bieten in Auktionen.

nen müssen. Ein Prosatext weist in der Regel keine annähernd exakten Strukturen auf. Dazu sind Grammatik und Semantik natürlicher Sprachen zu komplex, flexibel und mehrdeutig. Schwierigkeiten für die Analyse können entstehen, wenn Daten aus unterschiedlichen Quellen mit zwar vorhandenen, aber nicht übereinstimmenden Strukturmerkmalen auftauchen. Die Übergänge zwischen „strukturiert“ und „unstrukturiert“ sind hierbei fließend. Zwischen diesen Polen existiert eine beachtliche Zahl „semistrukturierter“ Daten in unterschiedlichen Graden der Strukturierung. Wie im Folgenden gezeigt werden wird, existiert für den Umgang mit unstrukturierten Daten eine Vielzahl von Verfahren.

Anforderungen an die Verfügbarkeit

Die Frage nach der Verfügbarkeit von Datenbeständen spielt eine entscheidende Rolle für die Verarbeitungsmöglichkeiten. Dabei sind zwei Herangehensweisen zu unterscheiden. Einerseits können Unternehmen, die mit Big Data arbeiten, von ihrem eigenen, bereits vorhandenen Datenbestand ausgehen und darauf aufbauend bei Bedarf weitere Datenbestände gezielt hinzukaufen, um neue Möglichkeiten der Auswertung und Nutzbarmachung der Daten zu finden oder um die Analysen zu präzisieren oder sonst zu verbessern. Andererseits kann ein Unternehmen auch zunächst festlegen, in welchen Bereichen es Erkenntnisse aus Datenbeständen gewinnen möchte, und daran anschließend nach Möglichkeiten suchen, diese Datenbestände aufzubauen.

Neben Start-ups, deren Geschäftsmodelle oft direkt auf der Analyse von Datenbeständen basieren, ist letztere Perspektive auch für solche Unternehmen von Interesse, die ihre strategische Geschäftsentwicklung und Akquise zunehmend daran ausrichten, welche Datenbestände hierdurch erschließbar werden. Die Relevanz dieses Geschäftsfelds wird nicht zuletzt an den starken Bewertungen und hohen Kaufpreisen von Firmen sichtbar, die solche Datenbestände besitzen oder jedenfalls potenziell große Datenmengen ansammeln werden. Das Beispiel Nest wurde in diesem Zusammenhang bereits genannt, gleiches gilt aber auch für Unternehmen wie Facebook oder Google selbst. Gerade die Geschäftstätigkeit des Letzteren ist seit Jahren darauf ausgerichtet, jenseits des Kern-

geschäfts der Internet-Suche neue Datenquellen über die Bereitstellung von Diensten wie E-Mail, Kartenservices oder einem Sozialen Netzwerk zu erschließen. Entscheidend für den Erfolg des Ansatzes ist der Aufbau riesiger Datenbestände aus unterschiedlichen Anwendungen, die nicht isoliert nebeneinander stehen bleiben, sondern anschließend miteinander verknüpft und korreliert werden können.

„Die neuen Big-Data-Datenmassen kommen hauptsächlich von privaten Unternehmen. In vielen Ländern der Welt ist die Kooperation zwischen dem öffentlichen Sektor und dem Privatsektor sehr gespalten und die Nutzung der Daten der Bevölkerung durch die Privatwirtschaft wird mit Skepsis betrachtet.“

Sabrina Juran, Technical Specialist,
Data and Research, United Nations
Populations Fund (UNFPA), Interview

Wahrscheinlichkeitsrechnung als Methode der Auswertung

Eine wichtige Grundlage zum Verständnis von Datenanalysen ist die Wahrscheinlichkeitsrechnung. Ein Grundprinzip aller Wahrscheinlichkeitsangaben ist, dass sie keine Aussagen über den Einzelfall zulassen, sondern nur Prognosen im Rahmen von Wahrscheinlichkeiten ermöglichen. Hierfür ist der Begriff der Korrelation entscheidend. Eine Korrelation beschreibt die Abhängigkeit einer Größe von einer anderen. Wenn die Korrelation in einem bestimmten Fall hoch ist, dann erlaubt das die Ableitung einer Aussage für einen Wert von einem anderen, wiederum im Rahmen der Wahrscheinlichkeiten.

So besteht beispielsweise eine beobachtbare Korrelation zwischen dem Konsum von Eiscreme und dem Verbrauch von Mineralwasser. Wenn festgestellt wird, dass die Umsatzzahlen von Eiscremeherstellern steigen, lässt sich mit einer gewissen Wahrscheinlichkeit die Aussage treffen, dass auch Mineralwasserhersteller mehr Umsatz machen werden.

Ein klassisches Beispiel für eine einfache Datenanalyse ist die Prozessoptimierung. In sich wiederholenden Prozessen können Daten über Produktions-

fehler erfasst werden. Findet sich in den Daten eine Korrelation der Produktionsfehler mit der Tageszeit, dann ergibt das einen Anhaltspunkt für eine Optimierung. So kann beispielsweise eine Korrelation darauf hindeuten, dass Mitarbeiter am Fließband im Idealfall nach drei Stunden eine Pause machen müssen oder dass ein Industrieroboter nach 48 Stunden gewartet und mit neuem Schmieröl versorgt werden muss.

Zunehmend kommen in der Datenanalyse selbstlernende Verfahren zum Einsatz, die aus der Forschung zur Künstlichen Intelligenz stammen. Mustererkennungsverfahren oder maschinelles Lernen mittels neuronaler Netze oder vereinfachter Ansätze ermöglichen das Auffinden von regelhaften Strukturen in Datenbeständen. Auch diese Verfahren fußen letztlich auf statistischen Berechnungen.

„Jeder, der nur auf die Technologie schaut, hat noch nicht begriffen, was die Automatisierung für die Gesellschaft bedeuten kann.“

Alexander Mankowsky, Zukunftsforscher bei der Daimler AG, öffentliche Veranstaltung Stuttgart, 25.06.2015

Korrelation und Kausalität

Die Analyse von Daten kann Erkenntnisse hervorbringen, die sich für Wirtschaft, Wissenschaft und politische Gestaltung nutzbar machen lassen. Korrelation kann aber nicht mit Kausalität gleichgesetzt werden. Nur weil eine Auswertung von Daten Abhängigkeiten zwischen zwei Größen zeigt, folgt daraus keineswegs, dass eine Größe die andere verursacht. Es ist ebenso möglich, dass es sich genau umgekehrt verhält oder dass beide Werte gleichermaßen von einer dritten Größe abhängen. Genauso könnte die scheinbare Korrelation auch reiner Zufall sein, und die beiden Größen sind in Wirklichkeit völlig unabhängig voneinander.

Um beim oben genannten Beispiel zu bleiben: Aus der Beobachtung, dass erhöhter Konsum von Speiseeis und erhöhter Verbrauch von Mineralwasser mit einer bestimmten Wahrscheinlichkeit gleichzeitig auftauchen, zwischen diesen Größen also eine Korrelation besteht, lässt sich nicht folgern, dass hier ein kausaler Zusammenhang vorliegt. Die Schlussfolge-

rung, dass der Verzehr von Eiscreme besonders durstig macht und deshalb der gestiegene Mineralwasserkonsum dadurch bedingt ist, dass die Leute mehr Eis gegessen haben, ist unzulässig. Überzeugender ist die Annahme, dass die beiden Werte von einem dritten abhängen, nämlich dem heißen Sommerwetter.

Fest steht lediglich anders herum, dass es ohne Korrelation keine Kausalität geben kann. Diese Erkenntnis ist keineswegs trivial, denn es macht die Falsifikation von Annahmen erst möglich. Wenn gezeigt werden kann, dass zwischen Ereignis A und Ereignis B entgegen der ursprünglichen Hypothese keine Korrelation besteht, dann ist damit auch bewiesen, dass es an einer Kausalität zwischen den Ereignissen fehlt.

Lässt sich also für einen bestimmten Zeitraum beobachten, dass der Konsum von Mineralwasser signifikant steigt, der Verzehr von Eiscreme aber sinkt, dann ist damit immerhin gezeigt, dass kein kausaler Zusammenhang zwischen den beiden Werten besteht.

Das Problem aktueller Big-Data-Analysen ist oft, dass sie so viele verschiedene Faktoren gleichzeitig in eine Korrelationsanalyse einbeziehen, dass einzelne Ursachenzusammenhänge nicht mehr erkennbar sind. Hier stößt die „Weisheit“ von Big Data an ihre Grenzen. Es könnte gezeigt werden, dass unter bestimmten Bedingungen mit einer hohen Wahrscheinlichkeit ein bestimmtes Ereignis eintritt. Warum das so ist, bleibt aber gänzlich unbekannt. Aus diesem Grund ist der aus Big-Data-Auswertungen gezogene Erkenntnisgewinn stets und inhärent limitiert.

Blinde Flecken

Vor der Digitalisierung waren Unternehmen und andere Akteure auf Umfragen und Zählungen angewiesen, um Datenmaterial für Analysen zu erhalten. In der umfassend vernetzten Gesellschaft hingegen entstehen die benötigten Daten nicht nur in wesentlich höherer Quantität, sondern oft auch in höherer Qualität – und das ohne aufwendige Erhebungen, sondern beinahe automatisch als Nebenprodukt von Primäranwendungen. Ein relevanter Teil traditioneller empirischer Marktforschung erfolgt auf Grundlage von Umfragen. So wird beispielsweise versucht, die Effektivität einer bestimmten Werbemaßnahme oder den Erfolg eines Produkts zu ermitteln, um bei Bedarf gegensteuern oder nachbessern zu können.

Die auf diese Weise gesammelten Aussagen sind aber nur dann aussagekräftig, wenn es gelingt, eine repräsentative Auswahl von Befragten zu treffen und zur Teilnahme zu bringen. Selbst dann bleibt das Risiko, dass die Aussagen zu einem gewissen Grad mit Ungenauigkeit behaftet sind. Verzerrungen können daraus resultieren, dass Befragte mitunter bewusst oder unbewusst falsche Angaben machen. Wenn ein Datenbestand alle (potenziellen) Kunden mit einbezieht und auf der unmittelbaren Beobachtung ihres Verhaltens aufbaut, dann sind die Daten von erheblich höherer Qualität. Ein offensichtliches Beispiel ist das Online-Marketing. Durch Tracking im Internet wird das Verhalten der Nutzer detailliert und umfassend aufgezeichnet. Auf diese Form des Marketings spezialisierte Firmen können zuordnen, welche Webseiten ein bestimmter Nutzer besucht hat. Die dabei hinterlassenen Datenspuren – zumeist über den Einsatz sogenannter Cookies – geben Aufschluss darüber, wer jeweils welchen Werbemitteln ausgesetzt war und in welchen Fällen diese in eine Kaufentscheidung mündeten.

Nicht vergessen werden sollte, dass jeder Datenbestand stets unvermeidliche blinde Flecken enthält. Prognoseaussagen können immer nur in Bezug auf die Grundgesamtheit getroffen werden. Eine Prognose, die auf der Beobachtung des Verhaltens von Nutzern eines bestimmten Sozialen Netzwerks basiert, ist selbst dann, wenn das Netzwerk sehr groß ist, nicht notwendigerweise auf die Gesamtgesellschaft übertragbar. Sehr große Datenbestände, wie sie für Big Data kennzeichnend sind, verleiten dabei leicht zu Verallgemeinerungen.

Wenn Entscheidungen mit gesamtgesellschaftlicher Tragweite auf solcherart vollzogenen Analysen basieren, besteht die Gefahr, dass das Ergebnis dieser Entscheidungsprozesse großen Teilen der Bevölkerung nicht gerecht wird und ihre Interessen und Bedürfnisse vernachlässigt werden. Ein Beispiel: Wenn die Analyse der Verkehrsströme in einer Großstadt anhand der Auswertung der Nutzungsdaten von Smartphones erfolgt, dann gehen in das Ergebnis nur die Verhaltensweisen und Präferenzen derjenigen Bewohner ein, die sich ein solches Gerät leisten können. Wird anschließend auf Grundlage der Analyse der Nahverkehr neu ausgerichtet, dann kann es deshalb sein, dass die Bedürfnisse sozial benachteiligter Bevölkerungsschichten nicht oder nur ungenügend Berücksichtigung finden, da von Mitgliedern dieser Gruppe keine Daten in die Berechnung eingeflossen sind.

„In der Diskussion um Big Data wird relativ häufig vorausgesetzt, dass durch die Datenanalyse ein Fortschritt erzielt wird, der nicht weiter hinterfragt wird. Es ist ein berechtigter Einwand zu fragen, ob sich bei den verschiedenen Big-Data-basierten Modellen das Ergebnis je Einzelfall auch verifizieren lässt.“

Christian Hawellek, wissenschaftlicher Mitarbeiter am Institut für Rechtsinformatik der Leibniz Universität Hannover,
1. Expertenworkshop, 26.02.2015

INTERVIEW MIT SILKE JANDT

Rechtskonforme Technik wäre ein großer Vorteil für alle

? Das Thema Big Data ist in aller Munde, und die damit verbundenen Chancen und Risiken werden kontrovers diskutiert. Bei provet befassen Sie sich unter anderem auch mit diesem Thema. Was ist provet und wie wirkt sich die Debatte auf Ihre Arbeit bei provet aus?

Silke Jandt: provet ist eine Forschungsgruppe an der Universität

Kassel, die sich grundsätzlich mit dem Thema verfassungsverträgliche Technikgestaltung befasst. In diesem Zusammenhang sind wir in viele interdisziplinäre Forschungsprojekte eingebunden, in denen Juristen mit Informatikern, Wirtschaftsinformatikern, Psychologen, Soziologen und weiteren Fachdisziplinen zusammenarbeiten. Wir versuchen, im interdisziplinären Dialog technikbezogene Fragestellungen wissenschaftlich zu beantworten.

Ich bin an vielen Projektanträgen direkt beteiligt, und man kann beobachten, dass das Thema Big Data in den letzten zwei Jahren vermehrt in den Forschungsanträgen vorkommt. Dabei geht es vor allem darum, wie man Big Data sinnvoll einsetzen kann. Da besteht noch viel technischer Forschungsbedarf. Seit einigen Jahren verlangen die Geldgeber bei solchen technikbezogenen Projekten, die durchaus große Chancen für die Gesellschaft bieten kön-

Dr. Silke Jandt

Dr. Silke Jandt ist seit Oktober 2015 Vertreterin des Lehrstuhls für Öffentliches Recht, Informationstechnologierecht und Rechtsinformatik an der Universität Passau. Sie ist als assoziierte Habilitandin Mitglied des DFG-Graduiertenkollegs „Privatheit und Vertrauen für mobile Nutzer“ der TU Darmstadt und

der Universität Kassel, das das Ziel verfolgt, neue Lösungen für den Schutz der Privatsphäre bei der mobilen Internet-Nutzung zu entwickeln. Von 2011 bis 2015 war sie Geschäftsführerin und seit 2008 wissenschaftliche Mitarbeiterin der Projektgruppe verfassungsverträgliche Technikgestaltung (provet) an der Universität

Kassel. In dieser Zeit hat sie in zahlreichen interdisziplinären Forschungsprojekten mit dem Ziel der datenschutzkonformen Technikgestaltung mitgearbeitet. Sie promovierte 2008 mit der Arbeit „Vertrauen im Mobile Commerce – Vorschläge für die rechtsverträgliche Gestaltung von Location Based Services“.

„Dafür versuchen wir, die unterschiedlichsten Sichtweisen zu berücksichtigen, die verschiedenen Interessen auszubalancieren.“

Silke Jandt

nen, von vornherein, dass Juristen beteiligt sind, die das Thema Datenschutzrecht behandeln. Damit soll sichergestellt werden, dass das, was technisch möglich ist, auch rechtlich so realisiert werden kann.

? Also die Juristen als Gewissen der Techniker?



Foto: Fotostudio Wilder, Göttingen

SJ: Ganz so weit würde ich nicht gehen. Wir wollen die Technik nicht ausbremsen, sondern sie so gestalten, dass sie gesellschaftsverträglich ist – und zwar für alle. Dafür versuchen wir, die unterschiedlichsten Sichtweisen zu berücksichtigen, die verschiedenen Interessen auszubalancieren und möglichst sofort in die Technik zu implementieren. Dahinter steht der Gedanke, dass für das, was technisch nicht realisiert wird, sich die Juristen keine Regeln mehr ausdenken müssen. Denn Regeln können gebrochen werden. Man muss überprüfen, ob Regeln eingehalten werden. Wenn die Technik von vornherein rechtskonform ist, ist das ein großer Vorteil für alle.

? Vor allem im medizinischen Bereich wurde schon lange mit großen Datenmengen gearbeitet. Was ist das Neue bei Big Data?

SJ: Dafür muss man das Schlagwort Big Data etwas genauer fassen. Oft wird von Big Data gesprochen, und jeder meint etwas anderes. Inzwischen werden die meisten Lebensbereiche durch elektronische Informations- und Kommunikationsmittel unterstützt. Dadurch sind mittlerweile unglaublich viele Daten digital vorhanden. Das ist die Grundsituation: diese große Datenmenge, die da ist. Bei Big Data geht es nun darum, diese Datenmengen sinnvoll auszuwerten, die Ergebnisse zu visualisieren, sodass Menschen damit etwas anfangen können, und daraus neue Erkenntnisse zu gewinnen.

? Welche neuen Erkenntnisse und Möglichkeiten bietet denn Big Data konkret? Können Sie Beispiele nennen?

SJ: Big Data wird zum Beispiel eingesetzt, um die Ausbreitung von Epidemien nachzuverfolgen.



RECHTSKONFORME TECHNIK WÄRE EIN GROSSER VORTEIL FÜR ALLE

“ Ich glaube nicht, dass man die Infrastruktur des Sammelns und die eigentlichen Daten voneinander trennen kann.

Silke Jandt



Man kann Gesundheitsdaten auswerten, um neue Erkenntnisse zu Heilungsprozessen zu bekommen. Oder im Verkehrsbereich kann man Big Data einsetzen, um Staus zu vermeiden und die Verkehrsflüsse zu optimieren.

? Wo sehen Sie die Risiken?

SJ: Der Fantasie sind keine Grenzen gesetzt, wo man Big Data einsetzen kann. Die Risiken sind also vielfältig. Man kann anhand von Big Data Prognoseentscheidungen treffen, wie wahrscheinlich es ist, dass jemand in Zukunft auf eine bestimmte Weise handeln wird. Das wird heute schon bei Kreditentscheidungen gemacht, kann aber auch im Gesundheitsbereich geschehen: Wie verhält sich jemand, wie sind seine Chancen, gesund zu bleiben,

wie wahrscheinlich ist es, dass er schwer krank wird? Im Marketing kann man feststellen, was eine Person zu einem bestimmten Zeitpunkt kaufen wird, und entsprechend Werbung schalten.

? Es werden normalerweise individuelle Daten von Personen gesammelt. Muss dieses Sammeln schon reguliert werden, oder reicht es, hinterher aufzupassen, wie die Daten zum Beispiel von Unternehmen verwendet werden?

SJ: Ich glaube nicht, dass man die Infrastruktur des Sammelns und die eigentlichen Daten voneinander trennen kann. Ich kann nachvollziehen, dass Informatiker sich das wünschen, aber es gibt auch nach bestehendem Datenschutz-

recht genug Möglichkeiten, sinnvoll zu forschen. Ich glaube, dass man im Gegenteil mehr technische Instrumente einsetzen sollte, um zu verhindern, dass ein Personenbezug entsteht.

In der Forschung braucht man in der Regel den Personenbezug gar nicht. Wenn Sie Gesundheitsdaten und Heilungsprozesse analysieren wollen, dann ist es völlig egal, ob es Frau Müller oder Herr Maier war, der oder die diese Daten geliefert hat. Entscheidend ist, dass es eine männliche Person in einem bestimmten Alter und mit einer bestimmten Vorgeschichte war. Informatiker scheuen sich manchmal, den Mehraufwand zu betreiben, um wirklich anonyme Datensätze zu erhalten. Hinterher Kontrollmaßnahmen einzuführen, ist mühsam und schwer zu kontrollieren.

? Eine Frage zu einem konkreten Beispiel: Mit vernetzten Autos, die die Fahrweise nachverfolgen, entstehen Daten, die Rückschlüsse über Personen zulassen: Ob man viel nachts unterwegs ist, ob man riskant fährt oder nicht. Wer soll wissen dürfen, was Sie wo und wann mit Ihrem Auto machen? Und ist es in Ordnung, wenn etwa Versicherungen aus diesen Fahrdaten Schlussfolgerungen ziehen, die sich auf den Vertrag auswirken?

SJ: Das smarte Auto ist datenschutzrechtlich schon problematisch. Es kommt aber darauf an, was Sie damit erreichen wollen. Wenn Sie in Richtung selbst fahrendes Auto denken, gibt es da viele Vorteile. Die bisherigen Tests deuten darauf hin, dass weniger Unfälle passieren, weil der Mensch als Fahrer fehleranfälliger ist als ein Computer. Die Daten, die dabei anfallen, braucht außer dem Fahrzeug und dem Fahrer niemand.

Anders sieht es aus, wenn wir über Versicherungsmodelle reden, die auf Tracking beruhen: Ich lasse mein Fahrverhalten kontrollieren, das wird an meine Versicherung

geschickt, und ich bekomme dann einen besseren Tarif. Aus rechtlicher Sicht kann man erst mal sagen, wenn derjenige das will und diesen finanziellen Vorteil nutzen möchte, ist das eine Form der informationellen Selbstbestimmung. Das ist ja die Grundlage des Datenschutzes. Allerdings ist es nicht immer sicher, ob das wirklich eine autonome Entscheidung ist. Wo sind die Grenzen, wo ist der finanzielle Druck so hoch, dass man seine Daten hergibt, obwohl man das vielleicht nicht möchte?

Wenn Sie die Daten einzelner Fahrzeuge nehmen, um den Verkehr optimaler zu lenken, dann müsste es eine zentrale Stelle geben, wo die Daten gesammelt und ausgewertet werden. Dafür braucht man nicht unbedingt zu wissen, wer welches Auto fährt. Das heißt, da könnte man die Daten anonymisieren. Auch wenn der Hersteller des Autos Daten sammelt, um zum Beispiel die Technik zu verbessern, braucht er keinen Personenbezug. Man muss also immer im Einzelfall schauen, wer bekommt welche Daten und zu welchen Zwecken.

? Wie sieht es aus, wenn wir das in den Gesundheitsbereich übertragen? Smart Watches, Fitnessarmbänder und Fitness-Tracker

sind auf dem Vormarsch. Ist es von Vorteil, wenn die Gesundheitsdaten von Nutzern bekannt sind?

SJ: Das ist ein sehr spannendes und kontroverses Thema. Wenn diese Fitnessanwendungen, Gesundheitsarmbänder oder Tracking-Apps dazu führen, dass Menschen sich gesünder verhalten, sportlicher werden und sich besser ernähren, ist das erst einmal positiv. Ich glaube allerdings, dass viele nicht wissen, was sie da tun und was mit ihren Daten passiert. Und das ist das Bedenkliche daran. In den wenigsten Fällen bleiben die Daten nur auf dem Endgerät des Nutzers. Sie landen im Regelfall auf dem Server des Unternehmens – wo auch immer es sitzt, wer auch immer das ist.

Wir haben in der Vergangenheit vom gläsernen Bürger gesprochen, aber heute hat das noch mal eine ganz andere Qualität, vor allem, wenn man sich Gesundheits- und Aktivitätsdaten anschaut. Es besteht die Gefahr, dass es in Zukunft heißt, ich habe kein Recht, ungesund zu leben. Ich muss mich gesund verhalten, richtig essen, Sport treiben. Wenn wir das als einzige Maxime ansehen, würde das unsere Welt sehr verändern.



RECHTSKONFORME TECHNIK WÄRE EIN GROSSER VORTEIL FÜR ALLE

“ Den Versicherungstarif davon abhängig zu machen, welche Gesundheitsdaten Menschen an die Krankenkasse übertragen, ist für mich ein Abschied von der Solidargemeinschaft.

Silke Jandt

➤ ? **Wie stehen Sie zu Krankenversicherungen, die Gesundheitsdaten von Fitness-Apps und Armbändern auswerten und in ihre Tarifmodelle einbeziehen?**

SJ: Den Versicherungstarif davon abhängig zu machen, welche Gesundheitsdaten Menschen an die Krankenkasse übertragen, ist für mich ein Abschied von der Solidargemeinschaft. Damit würde das Prinzip Krankenkasse zerstört werden. Das sehe ich als großes Problem an.

Das ist etwas anderes als die Autoversicherung, da Gesundheit viel elementarer ist als das Fahren eines Autos: Auf das Autofahren kann ich nämlich auch verzichten. Nicht alle Krankheiten können verhindert werden, indem ich mich gesund ernähre oder

Sport treibe – viele sind Veranlagung. Manche Menschen sind von vornherein aufgrund ihrer Gene benachteiligt. Die würden dann zusätzlich bestraft werden, wenn sie aufgrund ihrer Daten eine schlechtere oder teurere Krankenversicherung bekommen würden. Das ist keine Gesellschaft, die ich will.

? **Was wünschen Sie sich in der Debatte um die Ausgestaltung und Nutzung von Big Data? Brauchen wir mehr Geschwindigkeit, mehr Ruhe oder andere Blickwinkel?**

SJ: Es ist vor allem wichtig, die unterschiedlichen Interessen einzubeziehen. Man muss sich überlegen, was Big Data erreichen kann – nicht nur in tech-

nischer Hinsicht, sondern auch bezogen auf die gesellschaftliche Zielsetzung. Es gibt immer wieder Entwicklungen in der Forschung, dass Dinge möglich werden, aber wir als Gesellschaft uns dagegen entscheiden – das Klonen als Beispiel. Menschen dürfen nicht geklont werden, auch wenn es möglich wäre. Da gibt es eine ganz klare Entscheidung.

Wir sollten uns als Gesellschaft Gedanken machen, wo wir eine endgültige Grenze setzen möchten. Das mit den Daten wird immer weitergehen. In Zukunft wird es von jungen Menschen vom ersten Lebenstag an digitale Daten in irgendwelchen Datenbanken geben, sodass ihr ganzes Leben digital abgebildet wird. Und wenn alles zu allen Zwecken ausgewertet werden darf, dann haben wir ein ganz großes Problem. ▣

1.4 Anwendungsgebiete

Big-Data-Anwendungen kommen inzwischen in immer mehr Bereichen sowohl im privatwirtschaftlichen Sektor als auch in staatlichen Kontexten zum Einsatz.

Durch das Vorhandensein umfassender Datenbestände ergeben sich für staatliche Akteure neue Möglichkeiten, ihre Entscheidungen präziser an den gesellschaftlichen Bedürfnissen auszurichten. Allerdings werden diese Potenziale bislang nur zögerlich genutzt. Zwar greift beispielsweise die Stadtplanung schon lange auf empirische Erhebungen zurück, ein verstärkter Rückgriff auf Big Data aber könnte vollkommen neue Möglichkeiten eröffnen.

Ein weiterer Aspekt der Verwendung von Big-Data-Analysen für staatliche Zwecke ist die bereits erwähnte massenhafte geheimdienstliche Auswertung von Kommunikationsmetadaten, die unter anderem von den Vereinigten Staaten als Grundlage für Drohnenoperationen genutzt wird, um des Terrorismus Verdächtige gezielt zu töten. Diese „personalisierte Kriegführung“ (Klaus Mainzer) reklamiert, allein anhand bestimmter Kommunikationsmuster determinieren zu können, bei welchen Individuen es sich um Kämpfer handelt, die die Sicherheit der USA und ihrer Verbündeten gefährden, sodass aus der Perspektive der Gefahrenabwehr ihre präventive Tötung angezeigt und gerechtfertigt ist.

Auch in der innerstaatlichen Gefahrenabwehr durch Polizeibehörden kommt die Analyse großer Datenbestände immer häufiger zum Einsatz. Sogenanntes Predictive Policing, also voraussagende Polizeiarbeit, nutzt Big Data, um Verbrechenschwerpunkte in bestimmten Stadtteilen zu bestimmen oder die Tageszeiten, zu denen es vermehrt zu bestimmten Straftaten kommt, festzustellen, um dann mit vermehrten Streifen oder sogar ganz gezielten präventiven Einsätzen der Kriminalität entgegenzuwirken.

In der Wissenschaft ist die Nutzung von Big Data bereits seit Längerem etabliert und gehört in vielen Bereichen zum grundlegenden Rüstzeug. Selbstverständlich gibt es beträchtliche Abstufungen. Während in der experimentellen Elementarteilchenphysik, wie sie am Genfer CERN praktiziert wird, Big Data zum Kerngeschäft gehört, ist die Datenanalyse beispielsweise in der Rechtswis-

senschaft jenseits der Kriminalistik bisher kaum ein Thema. Dabei kann davon ausgegangen werden, dass auch in scheinbar datenfernen Disziplinen interessante Erkenntnisse durch die Nutzbarmachung von Datenanalysen gewonnen werden könnten. Erste Ansätze existieren bereits. So versucht das Projekt „Argumentum“, eine automatisierte Analyse von Argumentationsstrukturen in Gerichtsurteilen zu entwickeln. Daneben könnte zum Beispiel auch eine Auswertung von Gesetzesverweisen in Urteilen und anderen Rechtstexten Aufschluss darüber geben, welche Normen praktisch relevant sind, um darauf aufbauend den Rechtsbestand zu verschlanken.

Auch in der empirischen Sozialforschung bieten sich erhebliche Potenziale. Gerade hier wurden Daten bislang in erster Linie über Befragungen gewonnen. Dank Big Data kann diese Disziplin nun anhand von Beobachtungsdaten, wie sie bei der Internet-Nutzung entstehen, ganz neuen Wirkzusammenhängen nachgehen. Hier besteht allerdings noch das Problem, dass wenig Datenmaterial für wissenschaftliche Forschungszwecke zur Verfügung steht.

„Durch die Erfassung digitaler Gesundheitsdaten ist es möglich geworden, medizinisch-wissenschaftliche Fragestellungen schneller und effizienter zu erforschen und herkömmliche Therapieformen zu hinterfragen und zu optimieren. Dafür brauchen wir aber leistungsfähige und nicht nur komplexe Datenbanken. Wir sind heute weitaus besser und schneller in der Lage, Gesamtzusammenhänge zu erkennen und Therapieformen umzustellen.“

Prof. Dr. Karl Max Einhäupl,
Vorstandsvorsitzender der Charité –
Universitätsmedizin Berlin, Interview

Neben der staatlichen und wissenschaftlichen Anwendung wird Big Data immer mehr kommerziell

genutzt. Im Frühjahr 2015 brachte der Branchenverband Bitkom einen Leitfaden mit zahlreichen Beispielen aktueller Geschäftsmodelle heraus, die sich auf Big-Data-Anwendungen stützen.¹³ Ein Beispiel für den privatwirtschaftlichen Einsatz von Big-Data-Analysen sind automatisierte Übersetzungen. Hier haben sich lernende Mechanismen den Ansätzen als überlegen gezeigt, die versuchen, grammatikalische Regeln im Vorfeld nachzubilden. Nahezu alle automatischen Übersetzungsdienste bauen darauf auf, durch die Analyse eines größeren Textkorpus Regeln zu gewinnen, die sie dann auf neue Texte, deren Übersetzung noch unbekannt ist, anzuwenden. Dabei ist es für das System weitestgehend irrelevant, für welche Sprachen es trainiert wird. Entscheidend sind allein die Masse und Qualität der Datenbestände, aus denen „gelernt“ wird. Frühere Ansätze, die auf bestehenden grammatikalischen Regeln und Wörterbüchern aufbauten, waren deutlich weniger erfolgreich. Das gleiche Prinzip kommt inzwischen auch bei der Optimierung von Spracherkennung zum Einsatz.

„Um neue Potenziale durch Big Data auszuschöpfen, müssen Unternehmen heraus aus den klassischen Industriestrukturen und dem siloartigen Denken. Agilität und interdisziplinäre Zusammenarbeit etwa in Labs sind sinnvoll, um die Umsetzung von Geschäftsideen zu beschleunigen.“

Sven Löffler, Business Development Executive, Big Data & Data-Driven Business T-Systems International GmbH, Konsultation

Sogar im professionellen Sport wird schon seit einiger Zeit zunehmend auf Big-Data-Analysen zurückgegriffen. Gerade in Mannschaftssportarten wie Baseball oder Fußball dient die Auswertung immenser Datenbestände, die aus Spiel- und Spielerbeobachtungen gewonnen wurden, sowohl der Verbesserung

der strategischen als auch der taktischen Ausrichtung eines Teams.

„Kurzfristig benötigt die deutsche Wirtschaft dringend mehr Data Scientists. Der nahezu neue Berufsstand verbindet analytische Fähigkeiten mit technischem Know-how, Kompetenz im Datenschutz und einem Gespür für Märkte, Innovationen und Geschäftsmodelle. Dafür sollten auch an Universitäten interdisziplinäre Ausbildungskonzepte entwickelt werden.“

Dr. Dirk Hecker, Geschäftsführer Fraunhofer-Allianz Big Data und Abteilungsleiter Knowledge Discovery, Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS, Konsultation

Darüber hinaus ist Big Data auch aus individueller Perspektive zunehmend ein Thema. Die „Quantified Self“-Bewegung, die sich der quantifizierbaren Selbsterfassung verschrieben hat, findet immer größere Verbreitung. Hieraus ergeben sich vielfältige wirtschaftliche Betätigungsfelder für Unternehmen. So beginnen beispielsweise Krankenversicherungen ein Interesse an den so entstehenden Datenbeständen zu entwickeln.¹⁴ Zugleich entstehen immer weitere Produktinnovationen wie die kürzlich eingeführte Apple Watch, die immer mehr persönliche Daten des Trägers erfassen können. Es ist davon auszugehen, dass zukünftig eine Vielzahl von Informationen des Alltags über Sensoren erfasst werden wird. Dadurch werden detaillierte Rückschlüsse über Ernährungsgewohnheiten, Gesundheitsentwicklung oder sogar über Lernfortschritte in der Schule möglich.

Neben Smart Watches und Fitness-Trackern für Smart-Health-Anwendungen ist in jüngerer Zeit auch der Bereich Smart Mobility in den Fokus der Öffentlichkeit gerückt. Hier geht es einerseits um

¹³ Bitkom (Hg.), „Big Data und Geschäftsmodell-Innovationen in der Praxis: 40+ Beispiele“, Berlin 2015, online: <https://www.bitkom.org/Bitkom/Publikationen/Big-Data-und-Geschäftsmodell-Innovationen-in-der-Praxis-40-Beispiele.html>.

¹⁴ Siehe Matthew Allen, „Versicherungen und die ‚Big-Data-Revolution‘“, Swissinfo.ch, 22. April 2015, http://www.swissinfo.ch/ger/ohne-wearables-keine-krankenversicherung-_versicherungen-und-die-big-data-revolution-/41389092.

Geräte, die das Fahrverhalten von Autofahrern umfassend protokollieren. Sie können detaillierte Bewegungsprotokolle erstellen sowie Geschwindigkeit und sonstiges Verhalten hinter dem Steuer aufzeichnen. Auch hier haben Versicherungen bereits begonnen, auf der Grundlage solcher Daten teilnehmenden Kunden abhängig von ihrem Fahrstil günstigere Versicherungsprämien einzuräumen.¹⁵ Die Daten sind aber auch für andere Analysen interessant: So könnte beispielsweise die Diskussion über geschlechtsspezifische Unterschiede im Fahrverhalten auf der Basis konkreter Daten geführt werden.

Zur Smart Mobility gehört neben dem Aspekt individuellen Fahrverhaltens auch die Gewinnung großer Datenbestände zur Optimierung von Verkehrsströmen oder für die Verkehrlenkung. Diese Nutzung von Big Data ist häufig im Grenzbereich zwischen staatlicher und privatwirtschaftlicher Aktivität angesiedelt, da sie häufig in öffentlich finanzierten, aber von privaten Unternehmen durchgeführten Projekten erfolgt (PPP – Public Private Partnerships).

Smart Health und Smart Mobility bieten sich besonders für eine Fokussierung im Rahmen des Gesamtprojekts „Braucht Deutschland einen Digitalen Kodex?“ an. In beiden Bereichen hat die Technologie bereits einen erheblichen Verbreitungsgrad erlangt. Deshalb lassen sich anhand der beiden Themenfelder die Chancen und Herausforderungen von Big Data exemplarisch darstellen.

„Den Wunsch nach innerer Selbstreflexion gibt es schon sehr lange. Selbst Goethe hat sich quasi selbst vermessen und dokumentiert. Die neue Dimension bezieht sich auf die Vernetzung dieser Informationen durch neue technische Möglichkeiten und die Chancen und Risiken, die sich daraus ergeben.“

Dr. med. Peter Langkafel, Gründer und Geschäftsführer der Healthcubator GmbH, 2. Expertenworkshop, 15.06.2015

¹⁵ So plant beispielsweise Deutschlands größter Autoversicherer, die HUK-Coburg, entsprechende Tarife einzuführen: „HUK-Coburg plant Telematik-Tarife bei Autoversicherung“, Heise.de, 21. Mai 2015, online: <http://www.heise.de/autos/artikel/HUK-Coburg-plant-Telematik-Tarife-bei-Autoversicherung-2661601.html>.

INTERVIEW MIT SABRINA JURAN

Big Data für Armutsbekämpfung und nachhaltige Entwicklung

? Die gesellschaftlichen Chancen und Risiken von Big-Data-Anwendungen werden zunehmend diskutiert. Wie nehmen Sie diese Debatte in Ihrer Arbeit wahr? Gibt es ein verstärktes Problembewusstsein, was die Risiken betrifft?

Sabrina Juran: Bei uns ist es momentan ganz spannend. In der UN hat der Generalsekretär 2013 zur

sogenannten „Datenrevolution“ aufgerufen. Das ist ein übergreifendes Konzept innerhalb der internationalen Gemeinschaft, das darauf abzielt, Daten zu nutzen, um Entwicklungsthemen zu stärken, also vor allem nachhaltige Entwicklung und Armutsbekämpfung. Die nationalen statistischen Systeme sollen gestärkt und vernetzt werden, damit sie eine Planungs- und Evaluierungsgrundlage bieten, um die Länder nach vorne zu bringen.

Das Ganze hängt zusammen mit den Millenniums-Entwicklungszielen, die 2015 auslaufen. Seit etwa drei Jahren ist die internationale Gemeinschaft dabei, eine neue Post-2015-Entwicklungsagenda zu beschließen. Dafür ist es sehr wichtig, eine Datengrundlage zu haben, um die Entwicklungsziele zu beobachten und zu bewerten, aber auch, damit Länder ihre nationalen wie auch regionalen Interventionen planen können. Als die ursprünglichen

Sabrina Juran

Seit 2008 arbeitet Sabrina Juran als Technische Mitarbeiterin für Daten und Bevölkerungsanalyse beim Bevölkerungsfonds der Vereinten Nationen (UNFPA). Während ihrer Anstellung bei UNFPA war Sabrina Juran in allen Bereichen der Bevölkerungsforschung tätig. Ihre Arbeiten um-

spannen Gebiete der Soziologie, Demografie und Medizin, mit Forschungsschwerpunkten auf der wechselseitigen Beziehung von internationaler Migration und nachhaltiger Entwicklung, der demografischen Dividende, neuen Technologien für medizinische Anwendung und Innovation in statistischen Verfahren.

Foto: © UNFPA and Vilandez Photography



“ Die Hoffnung besteht, dass Big Data neue Informationen zu Themen liefern kann, über die wir traditionell keine oder sehr geringe Quellen haben. **Sabrina Juran**

Millenniums-Entwicklungsziele im Jahr 2000 verabschiedet wurden, spielten Daten noch keine so große Rolle. Außerdem wurden die Länder, die diese Daten produzieren mussten, nicht ausreichend miteinbezogen.

Die nächsten Entwicklungsziele, die hoffentlich im September dieses Jahres verabschiedet werden, sind noch einmal um-

fassender und weitreichender. Dafür muss eine Datengrundlage geschaffen werden. Auf der einen Seite sollen die bestehenden Datensysteme erweitert werden, zum Beispiel die nationalen Statistikämter ausgebaut und die Kompetenzen des dortigen Personals gestärkt werden. Auf der anderen Seite möchten wir aber auch neue Technologien

und neue Datenquellen nutzen. In diesem Kontext spielt Big Data eine Rolle.

? Welche Vorteile hat Big Data in Ihrer Arbeit?

SJ: Die Hoffnung besteht, dass Big Data neue Informationen zu Themen liefern kann, über die wir traditionell keine oder sehr geringe

In ihrer Arbeit bei UNFPA berät sie nationale Statistikämter in methodologischen und konzeptionellen Fragen rund um die Umsetzung nationaler Bevölkerungszählungen und die Nutzung traditioneller und nicht traditioneller Datenquellen wie auch in analytischen Aspekten bezüglich der Integration von Bevölkerungs-

themen in nationale Entwicklungspolitiken. Für Armut- und Bevölkerungsschätzungen hat Sabrina Juran mit Kollegen aus Akademia, Privatsektor und NGOs analytischen Nutzen aus Mobilfunkdaten, Satellitenbildern und Textdateien aus Social Media und Suchmaschinen gezogen. Innerhalb des Kontexts der Datenrevolution

schreibt sie wissenschaftlich und in Medienblogs zu Big Data für nachhaltige Entwicklung und medizinischen Fortschritt.

Ihr Studium der Sozialwissenschaften hat Sabrina Juran in Deutschland, Spanien und den USA absolviert. 2015 promovierte sie zum Dr. phil. an der Humboldt-Universität zu Berlin.



BIG DATA FÜR ARMUTSBEKÄMPFUNG UND NACHHALTIGE ENTWICKLUNG

» **Deshalb werden bei Volkszählungen auch so selten die Zensus-Mikrodaten herausgegeben, weil realistisch das Risiko besteht, dass die Privatsphäre von Individuen verletzt werden kann.** [Sabrina Juran](#)



Quellen haben. Es geht dabei nicht darum, dass Big Data die traditionellen Quellen wie Umfragen, Volkszählungen oder Register ersetzt, sondern dass diese Informationen angereichert werden können mit Daten, die man sonst nicht erhält. Allerdings gibt es natürlich immer ein Pro und ein Contra.

? Können Sie ein konkretes Beispiel nennen?

SJ: Ein Projekt, mit dem ich mich in den letzten Jahren beschäftigt habe, waren Bevölkerungs- und Armutsschätzungen. Konkrete Analysen zu Armut auf lokaler Ebene durchzuführen, ist deshalb unglaublich schwierig. Zunächst wollten wir direkt mit den sogenannten Call Detail Records von Mobilfunkfirmen arbeiten. Call Detail Records sind die vollständigen

anonymisierten Verbindungsdaten von Mobiltelefonen, also wer wann wo wie lange einen Anruf getätigt hat. Das hat sich allerdings als schwierig herausgestellt, da die Anbieter die Daten nicht herausgeben wollten – zu Recht. Denn diese Daten sind höchst sensibel, man kann mit ihnen sehr detaillierte individuelle Profile herstellen. Dazu kommt noch, dass in vielen Ländern Mobiltelefone nicht ausreichend verbreitet sind.

In der Elfenbeinküste haben wir mit dem Mobilfunkunternehmen Orange zusammengearbeitet. Das hatte eine Data for Development Challenge gestartet und über eine gewisse Zeit seine Daten anonymisiert für Analysezwecke freigegeben. Wir konnten, basierend auf einer Umfrage aus der Elfenbeinküste, die zugänglich war, Indikatoren ableiten, die uns bei unserer Analyse in Afghanistan halfen.

Bei einem anderen Projekt haben wir versucht, Google-Suchbegriffe dazu zu nutzen, um Migrationsstatistiken zu erweitern. Migrationsstatistiken sind generell sehr unzuverlässig, da nur wenige Länder direkte Informationen dazu haben. In der Regel messen die Länder nicht, wie viele Menschen ausgewandert sind, sondern schauen, wie viele Menschen aus Land Y in Land X leben, und ziehen daraus ihre Rückschlüsse. Zu den monatlichen oder jährlichen Migrationsflüssen zwischen den Ländern wissen wir noch weniger. Man sollte denken, das wäre heute ganz einfach, weil wir ja meistens mit Flugzeugen reisen und am Flughafen registriert werden. Das stimmt aber nicht. Deshalb haben wir geschaut, wie weit man mit Ersatz-Indikatoren kommt, um Migration zu messen. Wir haben das in Australien getan, weil es

eines der wenigen Länder ist, die monatliche Migrationsflüsse messen und veröffentlichen. Deshalb konnten wir dort testen, inwieweit bestimmte Google-Suchen mit den tatsächlichen Flüssen von Migration korrelieren. Im Ergebnis war die Korrelation sehr stark. Es geschah zwar ein bisschen zeitverzögert, aber man konnte deutlich sehen, dass gewisse Suchbegriffe wie „jobs in Melbourne“ oder „moving with babies“ einen großen Zusammenhang aufwiesen mit den Migrationsquoten.

? Welche Risiken gibt es bei dem Einsatz von Big Data? Wo liegen die Probleme?

SJ: Ein großes Problem ist das der Rückführbarkeit auf eine individuelle Person. In einem Artikel „Unique in the Crowd“ von Kollegen des MIT Media Lab, den ich kürzlich gelesen habe, wurde erläutert, wie viele Datenpunkte tatsächlich notwendig sind, um Rückschlüsse auf ein Individuum zu ziehen. Erschreckenderweise sind es nur vier. Wenn die Daten fehlerbehaftet sind und Details geändert wurden, braucht man eventuell sieben Datenpunkte. Aber bei einer sauberen Analyse der Daten, zum Beispiel, wo sich das Handy einloggt hat, welche Webseite man sich angesehen hat

und Ähnliches, waren es gerade mal vier Datenpunkte, die es einem erlauben, Rückschlüsse auf ein Individuum zu ziehen. Das ist natürlich innerhalb der UN und der internationalen Gemeinschaft eine große Sorge, dass einzelne Personen so einfach identifiziert werden können. Deshalb werden bei Volkszählungen auch so selten die Zensus-Mikrodaten herausgegeben, weil realistisch das Risiko besteht, dass die Privatsphäre von Individuen verletzt werden kann.

Ein anderes Problem ist, dass vor allem digitale Datenquellen verzerrt sind. Wenn wir uns die digitalen Realitäten der Mitgliedsstaaten der UN anschauen, so sehen wir, dass die Internet-Verbreitung in den verschiedenen Ländern sehr unterschiedlich ist. Auch innerhalb der Länder gibt es große Ungleichheit, wer Zugang hat und wer nicht. Die Leute, die eigentlich am meisten Beachtung brauchen, tauchen in Big-Data-Analysen vielleicht gar nicht auf, weil sie zum Beispiel keine Smartphones haben, die Daten generieren. Ein ganz pauschales Beispiel: Oft ist es so, dass junge Frauen in ländlichen Regionen deutlich weniger Zugang zu digitalen Technologien haben als höher gebildete Männer in Städten. Die Schere ist unglaublich groß – und die Gefahr besteht, dass diejenigen, die so-

wieso schon ausgegrenzt werden, noch einmal hinten herunterfallen.

Oft werden die betroffenen Länder, in denen wir Daten sammeln und analysieren, nicht ausreichend in die Arbeit einbezogen. Das liegt auch daran, dass in vielen Ländern die Kapazitäten und Kompetenzen, die Daten zu verstehen, gar nicht oder nicht ausreichend vorhanden sind. Wir haben in vielen Ländern große Probleme, grundlegende Analysen durchzuführen. Wenn jetzt noch Big Data dazukommt, kann dies zu einer Überforderung oder Ablenkung führen. In diesem Fall kommen oft Drittparteien – in der Regel private Unternehmen – ins Spiel, anstatt die Leute vor Ort zu schulen und ihnen die Möglichkeit zu geben, ihre eigenen Daten zu analysieren. Dadurch zieht man die Verantwortung und den Sachverstand der Länder nach unten.

Zusätzlich handelt man sich ein Datenschutzproblem ein – denn in vielen Ländern wird die Zusammenarbeit mit privaten Unternehmen sehr kritisch gesehen. Auch in Deutschland ist das ja so.

? Wie groß ist das Risiko für Einzelpersonen, dass Profile über sie erstellt werden, egal ob von Regierungen oder von Privatunternehmen?



BIG DATA FÜR ARMUTSBEKÄMPFUNG UND NACHHALTIGE ENTWICKLUNG

“ Wir brauchen ein internationales Basisabkommen, nach dem sich die nationalen Regierungen richten können, wenn Daten produziert, publiziert und genutzt werden.

Sabrina Juran



SJ: Die verschiedenen UN-Organisationen – und auch wir – erstellen keine individuellen Profile. Unsere Datenanalysen finden auf regionaler und nationaler Ebene statt. Wir versuchen zwar, möglichst detaillierte Daten zu erhalten, aber die niedrigste Ebene ist die Kommunalebene. Wenn wir Daten über eine Stadt oder ein Dorf bekommen, sind wir schon glücklich. Bei allen Daten, die wir erhalten, zum Beispiel aus Volkszählungen, achten wir immer darauf, dass man keine Rückschlüsse auf das Individuum ziehen kann. Dafür gibt es Techniken, dass zum Beispiel die individuellen Daten verschoben werden.

Wir sind auch sehr vorsichtig, was die Veröffentlichung der Rohdaten angeht: Vor allem in kleinen Ländern der Welt lassen sich Individuen schnell identifizieren – wo diese Person lebt und arbeitet und

so weiter. Risiken gehen auch vom Privatsektor aus, von gewissen Regierungen oder von politischen Gruppen im Land, die Interesse am Zugang haben. Bestimmte Daten müssen verstärkt geschützt werden – ein Beispiel sind die schon erwähnten Call Detail Records, die Verbindungsdaten von Mobiltelefonen. Dazu wurde kürzlich ein Kodex veröffentlicht, an dem der UNFPA mitgearbeitet hat und der das Ergebnis hat, dass diese Daten nicht aus der Hand der Telefonanbieter gegeben werden dürfen, sondern die Analyse nur von den Telefonanbietern selbst oder firmeneigenen Forschungsinstituten durchgeführt werden dürfe.

? Auf welcher Ebene müssen denn solche Dinge geregelt werden? Durch Selbstverpflichtungen der Unternehmen wie bei dem

genannten Kodex? Auf gesetzlicher Ebene – national oder international?

SJ: Beides ist wichtig – sowohl auf nationaler als auch auf internationaler Ebene. Durch den internationalen Charakter des Internets und der Arbeit auf diesem Gebiet ist eine rein nationale Gesetzgebung zur Nutzung von Daten schwierig. Selbst innerhalb der Europäischen Union sind die Gesetzgebungen zur Pressefreiheit, zur Meinungsfreiheit und zum Datenschutz noch unterschiedlich. Wir brauchen ein internationales Basisabkommen, nach dem sich die nationalen Regierungen richten können, wenn Daten produziert, publiziert und genutzt werden. Gerade die großen Unternehmen wie Google agieren ja international, dagegen können Nationalstaaten nur angehen, wenn sie sich koordinieren. □

1.5 Big Data im Gesamtprojekt „Braucht Deutschland einen Digitalen Kodex“?

Wie die vorangegangenen Abschnitte bereits deutlich gemacht haben, handelt es sich bei Big Data um eine technologische Entwicklung, die weite Teile der modernen Gesellschaft in tief greifender Weise erfasst. Eine solche Entwicklung wird nicht ohne Verwerfungen vonstattengehen. Überall dort, wo Probleme und Interessenkonflikte zutage treten, bedarf es einer gesellschaftlichen Auseinandersetzung und im Ergebnis Regeln.

„Wir stehen bei Big Data immer noch am Anfang der Entwicklung. In Deutschland gibt es zwar eine starke Debatte diesbezüglich, aber uns fehlt ein ethisches und rechtliches Rahmenwerk zur Nutzung persönlicher Daten.“

Dr. David Deißner, Director Strategy and Programmes des Vodafone Instituts für Gesellschaft und Kommunikation,
2. Expertenworkshop, 15.06.2015

Regeln für Prozesse der Datenerhebung und -verarbeitung stellt traditionell vor allem das Datenschutzrecht auf. Grundsätzlich gestattet es die Datenverarbeitung personenbezogener Daten nur, wenn hierfür eine Einwilligung vorliegt. Diese kann sich aus dem Gesetz ergeben oder als explizite Einwilligungserklärung des Daten-Subjekts (in der Regel: einer Person) eingeholt werden. Die Methoden und Ansätze von Big Data allerdings zeigen die Grenzen dieses traditionellen Regelungsmodells auf. So bezieht sich der Anwendungsbereich des Datenschutzrechts lediglich auf die Verarbeitung von personenbezogenen Daten. Viele im Rahmen von Big-Data-Anwendungen gesammelte und verarbeitete Daten haben aber für sich genommen keinen Personenbezug im herkömmlichen Sinne. Um statistisch untermauerte Aussagen über Menschengruppen mit bestimmten Eigenschaften zu machen, sind nicht zwingend Daten notwendig, die einen Bezug zu einer Person aufweisen.

Auch in anderer Hinsicht stoßen datenschutzrechtliche Prinzipien an ihre Grenzen. Dies gilt zum Beispiel für die Grundsätze der Zweckbindung – dass der Zweck der Datenerhebung vor Beginn bestimmt sein muss – und der Datensparsamkeit. Bei Big Data geht es aber auch darum, möglichst viele Daten „auf Vorrat“ zu sammeln, um anschließend mithilfe von algorithmischen Berechnungen Erkenntnisse aus dem Bestand zu ziehen. Dabei ist erwünscht, dass überraschende Zusammenhänge zutage treten, die nicht vorhersehbar sind und damit mit dem ursprünglichen Zweck – so überhaupt einer formuliert worden ist – gar nichts zu tun haben. Welche Folgen die Analysemöglichkeiten haben, die aus einer immer tiefer gehenden Digitalisierung und damit Verdattung des Alltags resultieren, ist noch nicht abschätzbar. Der Siegeszug von Big Data wirft deshalb im Hinblick auf das überkommene Datenschutzrecht viele grundlegende Fragen auf.

„Es wäre wünschenswert, wenn die gesetzlichen Rahmenbedingungen in Bezug auf Big Data klar und verbindlich wären, sodass Unternehmen nicht dadurch gehemmt werden, dass sie in gesetzlichen und ethischen Grauzonen handeln.“

Dr. Dirk Hecker, Geschäftsführer Fraunhofer-Allianz Big Data und Abteilungsleiter Knowledge Discovery, Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS, Konsultation

„Wir brauchen ein neues Datenschutzrecht mit genaueren Definitionen. Das Datenschutzrecht stammt aus einer Zeit, in der wir überhaupt nicht wussten, wie die digitale Welt aussehen würde. Jetzt haben wir zumindest eine Ahnung. Es muss viel genauer erfasst werden, was erlaubt ist und was nicht.“

Dr. Daniela Mielchen, Rechtsanwältin bei Mielchen & Coll. Rechtsanwälte für Verkehrsrecht, Konsultation

Da die Einsatzgebiete von Big Data so unterschiedlich sind, erscheint es angezeigt, sich konkret auf bestimmte Bereiche zu beschränken. Hier bieten sich Smart Mobility und Smart Health als exemplarisch an, weil diese Anwendungsfelder mit ihren vielen, erst in jüngerer Zeit auf den Markt gebrachten technischen Neuerungen zwar begonnen haben, sich zu etablieren, aber dennoch bislang nicht Gegenstand einer umfassenden Diskussion über Regelungsdefizite und -bedürfnisse waren. Die gesellschaftliche Auseinandersetzung verspricht erheblichen Erkenntnisgewinn über potenzielle Konfliktslagen.

„Durch Big Data können vor allem dann gesellschaftliche Konflikte entstehen, wenn die Gesellschaft nicht diskutiert, welche Formen der Nutzung persönlicher Daten wünschenswert sind und welche nicht – und die Grenzen hier absteckt. Dieses Risiko erhöht sich, je mehr man sich den Trends verweigert und sich nicht damit auseinandersetzt.“

Florian Schumacher, Gründer Quantified Self Deutschland und Digital Health Consultant iic-solutions, Konsultation

Als übergreifendes Thema von Smart Mobility und Smart Health liegt der Fokus des Projekts auf Tracking-Technologien. Diese werden in beiden Bereichen zunehmend dafür eingesetzt, um Einzeldaten zu sammeln, die anschließend zu Big Data aggregiert und für Verarbeitung und Auswertung bereitgestellt werden. Anhand der drei Beispiele telematikbasierte Autoversicherung, Verkehrslenkung und Wearables wie Fitnessarmbänder oder Smart Watches, die allesamt auf Tracking zur Generierung der für ihre jeweiligen Zwecke erforderlichen Datenbasis zurückgreifen, werden die Chancen und Risiken der Technologie identifiziert. Hieraus können Erkenntnisse für einige maßgebliche Fragen abgeleitet werden: Wo gibt es Konfliktpotenzial, wo verlaufen die Konfliktlinien, und welche gesellschaftlichen Herausforderungen müssen zur Lösung der Konflikte bewältigt werden? Ziel

ist dabei stets, die Rolle herauszuarbeiten, die ein Digitaler Kodex für den Aushandlungsprozess spielen könnte, und welche Elemente er dazu notwendigerweise aufnehmen müsste. Anhand von Tracking kann der Grundkonflikt zwischen den Vorteilen der Nutzung von Bewegungs- und Verhaltensdaten und den Beeinträchtigungen der Privatsphäre und des Persönlichkeitsrechts der Nutzer besonders anschaulich herausgearbeitet werden.

Bei der Nutzung von Big Data für Smart Mobility und Smart Health kommt darüber hinaus der sogenannten Algorithmenethik besondere Relevanz zu, auf die deshalb noch gesondert eingegangen wird. Die jeweiligen Algorithmen, die der Datenanalyse und -auswertung zugrunde liegen, sind häufig ein Entscheidungskriterium für die Verfügbarkeit oder Ausgestaltung eines Produkts – zum Beispiel, ob eine Person eine bestimmte Krankenversicherung bekommt, wenn die Datenauswertung eine erhöhte Wahrscheinlichkeit für gewisse Krankheiten ergeben hat. Sie wirken sich so unmittelbar auf die Interessen und Bedürfnisse der betroffenen Individuen aus. Deshalb werden bereits seit einigen Jahren Forderungen nach einer handlungsleitenden Ethik für diesen Bereich laut. So erklärte der Internet-Unternehmer Stephan Noller bereits im Oktober 2012 in einem Beitrag für die Frankfurter Allgemeine Zeitung, es müsse eine Diskussion über Mechanismen zur Kontrolle von Algorithmen geben.

1.6 Der Projektverlauf

Das Teilprojekt „Big Data“ begann mit der Zusammenstellung einer hochrangig besetzten Expertengruppe, die sich erstmals im Februar 2015 traf, um sich dem Thema anzunähern und die Erkenntnisse und Thesen des ersten für das Projekt verfassten Themenpapiers zu erörtern.¹⁶ Es ging bei dem Workshop in erster Linie darum, den Rahmen des sehr breiten Themas „Big Data“ abzustecken und für den weiteren Projektverlauf geeignete Zuspitzungen zu identifizieren. Zu Beginn wurde der Themenkomplex umrissen, indem die wachsende Bedeutung von Big-Data-Anwendungen für unser tägliches Leben

¹⁶ Jan Schallaböck, „Big Data“, https://www.divsi.de/wp-content/uploads/2015/01/Themenpapier_Big-Data.pdf.

dargestellt wurde. Als Ausgangspunkt für die folgende Diskussion diente die Feststellung, dass Big Data für sich genommen weder gut noch schlecht ist. Entscheidend sei allein, wie man mit den Ergebnissen der Analysen und Auswertungen umgeht. Für die Erörterung der Steuerung und Regulierung, die notwendig sein könnte, sollten die folgenden Fragen als Orientierung dienen:

- Was ist die neue Qualität von Big Data im Vergleich zu Datensammlungen im vordigitalen Zeitalter?
- Welches sind die Merkmale, die durch die Digitalisierung neu hinzugekommen sind?
- Welchen gesellschaftlichen Handlungs- und Regulierungsbedarf gibt es im Bereich Big Data?

Nach einer allgemeinen Problembeschreibung einigten sich die Teilnehmer im weiteren Verlauf darauf, sich für das Projekt auf die genannten Teilbereiche Smart Mobility und Smart Health zu fokussieren. Als Grundlage sollten „ausgehandelte Geschichten“ dienen, die die Potenziale und Herausforderungen der Technologie für den jeweiligen Bereich deutlich machen. Ein Ergebnis des Workshops war, dass Big Data sowohl für den Gesundheits- als auch den Mobilitätssektor viele Chancen der Ressourcenoptimierung und Verbesserung verspreche, zugleich aber Risiken für Individuen und für bestimmte gesellschaftliche Randgruppen berge. Die Expertenrunde schloss mit dem Vorschlag, das Thema für den weiteren Projektverlauf

weiter zu präzisieren, um Bedürfnisse und Möglichkeiten der Regulierung identifizieren zu können.

Die Schlussfolgerungen der Expertenrunde dienen im Anschluss als Grundlage für die Erarbeitung eines weiteren Themenpapiers, das die beiden gewählten Teilbereiche Smart Health und Smart Mobility im Hinblick auf Big Data und die Technologie des Trackings zur Datengewinnung näher beleuchtet.¹⁷ Das Papier stellt Big-Data-Anwendungen in den beiden Feldern vor und benennt die Chancen und Herausforderungen, die sich aus den Entwicklungen für die Gesellschaft ergeben. Es liegt dem zweiten und dritten Kapitel dieses Berichts zugrunde.

Mitte Juni folgte die zweite Expertenrunde, zu der erstmals auch vier Vertreter von Unternehmen – jeweils zwei für jedes der beiden Unterthemen – eingeladen waren, um die Diskussion aus der Perspektive von Akteuren anzureichern. Die zwei Blöcke zu Smart Health und Smart Mobility begannen dabei jeweils mit einführenden fiktiven Geschichten, die konkrete Anwendungsszenarien in den Bereichen Gesundheit und Mobilität aufzeigten. Diese Geschichten eröffnen in diesem Bericht auch die beiden Kapitel zu den Themen. Im Verlauf der Debatte kristallisierten sich eine Reihe von Fragen heraus, die im Hinblick auf die Möglichkeit der Regulierung von Big-Data-Anwendungen – im Sinne des Gesamtprojekts eventuell durch Digitale Kodizes – zu beantworten wären.

¹⁷ Henning Lahmann, „Gesellschaftliche Konfliktlinien im Kontext von Big Data am Beispiel von Smart Health und Smart Mobility“, DIVSI 2015, https://www.divsi.de/wp-content/uploads/2015/09/Themenpapier_Konfliktlinien-Big_Data.pdf.

DIE EXPERTENGRUPPE

Das Projekt wurde von einer interdisziplinär ausgerichteten Runde hochrangiger Expertinnen und Experten aus Wissenschaft, Wirtschaft und Zivilgesellschaft beraten. Die Gruppe hat als Sounding Board das Projekt im Rahmen von Workshops bei der Erarbeitung der Ergebnisse unterstützt.



Foto: Katrin Binner, www.katrinbinner.de

Prof. Dr. Johannes Buchmann

Vizedirektor von CASED, Sprecher des Forschungszentrums CYSEC und des Sonderforschungsbereichs CROSSING und Professor für Informatik und Mathematik an der TU Darmstadt

„Heute und in Zukunft stehen uns gigantische Datenmengen aus allen Bereichen zur Verfügung. Ihre Analyse unterstützt das Verständnis für Zusammenhänge und erlaubt sehr genaue Prognosen. Diese Technologie hat ein enormes Potenzial. Viele fürchten sich davor, dass sie dadurch

gläsern werden. Andere verbinden damit große Hoffnungen, zum Beispiel für die Gesundheit der Menschen. Ich nehme an der Expertengruppe teil, um zu einem Diskurs beizutragen, der dazu führt, dass diese Technologie zum Wohl der Menschen eingesetzt wird.“



Foto: Berliner Beauftragter für Datenschutz und Informationsfreiheit

Dr. Alexander Dix, LL.M.

Berliner Beauftragter für Datenschutz und Informationsfreiheit

„Hinter dem Begriff ‚Big Data‘ verbergen sich sehr unterschiedliche Anwendungen, auf die das Datenschutzrecht differenzierte Antworten finden muss. Während die medizinische Forschung, ausgehend von Korrelationen, nach Kausalzusammenhängen sucht, beschränken sich kommerzielle Anwender häufig auf die Ermittlung von Korrelationen und ziehen daraus negative Konsequenzen für Einzelne und ganze Personengruppen. Auf die Probleme der informationellen Diskriminierung

und des Social Sorting hat das Datenschutzrecht bisher keine befriedigenden Antworten. Das muss sich ändern. Ein Digitaler Kodex könnte dazu beitragen.“

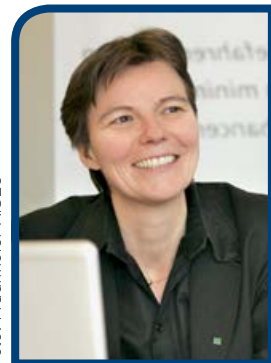
Prof. Dr. Claudia Eckert

Leiterin des Fraunhofer-Instituts für Angewandte und Integrierte Sicherheit (AISEC) und Professorin für IT-Sicherheit an der Technischen Universität München

„Big Data eröffnet sehr viele Chancen für Unternehmen, neue Dienstleistungen anzubieten. Beispielsweise können unterschiedlichste Daten über Maschinen und Anlagen erhoben und zusammengeführt werden, um frühzeitig mögliche Probleme zu erkennen und zu beheben. Auch für Verbraucher kann die Zusammenführung von Daten aus den verschiedensten Quellen und deren intelligente Auswertung viele Vorteile bieten – zum Beispiel in der Medizin oder beim Autofahren. Daten sind Werte und das ‚Öl‘ der Wissensgesellschaft.

Diese Werte zu schützen und dafür Sorge zu tragen, dass es nicht zu Datenmissbrauch und unerwünschter Datenweitergabe kommt, ist eine enorme Herausforderung. Es ist deshalb sehr wichtig, dass man frühzeitig mit allen gesellschaftlich Beteiligten in den Diskurs eintritt und die Chancen, aber auch die Risiken dieser Technologien offen debattiert und sich gemeinsam über Richtlinien und Prinzipien im Umgang mit diesem wichtigen Gut – also unseren Daten – verständigt. Das Projekt leistet hierzu einen sehr wertvollen Beitrag.“

Foto: Fraunhofer AISEC



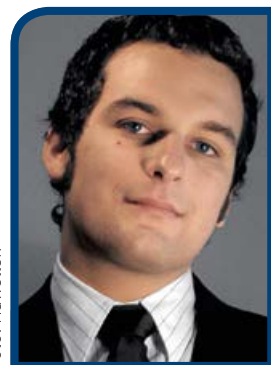
Christian Hawellek

Wissenschaftlicher Mitarbeiter am Institut für Rechtsinformatik der Leibniz Universität Hannover

„Big-Data-Applikationen sind datenschutzrechtlich schon deswegen sehr bedeutsam, weil sie sich hinsichtlich ihres Gefahrenpotenzials für Grundrechtspositionen schwer bis gar nicht adäquat mit den gegebenen rechtlichen Instrumenten erfassen lassen. Gleichzeitig sind die in der Analyse und Auswertung von Big Data ablaufenden Prozesse so komplex, dass die Integrität der erzielten Ergebnisse durch Menschen kaum noch nachvollziehbar sein dürfte. So

droht eine gefährliche Nähe zu den verbotenen automatisierten Einzelentscheidungen zu entstehen. Aus diesem Grund liegt es nahe, jenseits des bestehenden Rechtsrahmens über grundsätzlich neue legislative Ansätze nachzudenken, die sich – wo notwendig – auch von tradierten datenschutzrechtlichen Dogmen werden lösen müssen. Gleichzeitig gilt es, dabei auch die europäischen Entwicklungen in diesem Feld sorgfältig im Auge zu behalten.“

Foto: Hawellek



DIE EXPERTENGRUPPE



Foto: Steffen Leidel

Lorena Jaume-Palasi

Direktorin für Kommunikation und Jugendbeteiligung am Europäischen Dialog für Internet Governance (EuroDIG) und Dozentin

„Weder Daten noch Algorithmen tragen eine moralische Verantwortung. Für die Verarbeitung kleiner Datenmengen gelten dieselben ethischen Prinzipien wie bei der Verarbeitung großer, heterogener Datenmengen. Die potenziellen ethischen Risiken von Big Data verlangen nach keinem neuen ethischen Prinzip. Jedoch führen herkömmliche Erkenntnismethoden, wie das Prinzip der Induktion, das wir in der Mathematik seit Langem kennen, durch digitale

Induktionsverfahren wie Big Data zu erheblichen ethischen Unsicherheiten. Diese bedürfen einer Klärung durch eine breite ethische Debatte.“



Foto: Andre Kowalski

Stephan Noller

Gründer und Geschäftsführer der ubirch GmbH

„Wesentliche Bereiche unserer Gesellschaft werden schon in wenigen Jahren von Big-Data-Technologie geprägt sein. Egal, ob es automatische Übersetzungen, selbst fahrende Autos oder individuell zugeschnittene journalistische Angebote sind – wir werden von Algorithmen und Daten umgeben sein. Darin liegt viel Gutes, aber natürlich auch Gefahren. Die Idee eines ‚Digitalen Kodex‘, um diese für

den Gesetzgeber häufig zu schnellen Entwicklungen wenigstens ein bisschen normativ zu begleiten, halte ich für absolut richtig und wegweisend in der Debatte.“

Dr. Bernhard Rohleder

Hauptgeschäftsführer Bitkom e.V.

„Wie schaffen wir die Balance zwischen Datensparsamkeit und Datenvielfalt? Wie nutzen wir Datenschätze und sichern gleichzeitig den Datenschutz? Wie schöpfen wir die Potenziale digitaler Technologien voll aus, verhindern aber ihren Missbrauch? Wie entwickeln wir eine leistungsfähige Datenwirtschaft am Standort Deutschland und berücksichtigen dabei die besonderen Schutzbedürfnisse unserer Gesellschaft? Wie bringen wir die Rechte

auf informationelle Selbstbestimmung und Informationsfreiheit in Einklang? Wie wägen wir die informationelle Selbstbestimmung gegenüber anderen Grundrechten wie jenem der körperlichen Unversehrtheit ab? Wie lässt sich die breite gesellschaftliche Debatte um den Datenschutz in Schwung halten und gleichzeitig versachlichen? Diese Fragen sind auch im weltweiten Maßstab unbeantwortet. Nicht nur Deutschland braucht einen Digitalen Kodex.“

Foto: Bitkom



Tobias Schwarz

Leitender Chefredakteur Netzpiloten.de

„Big Data hat einen enormen Einfluss auf unsere Gesellschaft. Es verändert das persönliche Verhalten jedes Einzelnen, nach welchen Kriterien wir Entscheidungen treffen, wie wir Unternehmen führen oder wonach wir forschen. Der Um-

gang mit den Daten erfordert vollkommen neue Kulturtechniken von uns, denn noch nie war die Menge an zugänglichen Informationen so groß, es war aber auch noch nie so schwer, sich auf das Wesentliche zu fokussieren.“

Foto: Kai Oliver Goldmann, Shootbook



Prof. Dr. Sebastian Spaeth

Professor für Betriebswirtschaftslehre und Digitale Märkte an der Universität Hamburg

„Aus mehreren, jede für sich harmlos erscheinenden Datenquellen lassen sich mittels Big-Data-Methoden erstaunliche, interessante, aber auch erschreckend detaillierte Schlussfolgerungen ziehen. Das ist für Firmen und andere Organisationen wertvoll und für Kunden oft nutzbringend. Allerdings lassen sich diese Da-

ten auch missbrauchen und dringen tief in die Privatsphäre der Betroffenen ein. Wir als Gesellschaft müssen uns überlegen, ob und welche Regeln, Gesetze oder Normen notwendig, nützlich und gewünscht sind, um den Umgang mit dieser Datenflut und den daraus gewonnenen Informationen zu regeln und zu steuern.“

Foto: Manuel Fischer, www.freshpixel.ch



Smart Mobility

- Wer trägt die Verantwortung beim automatisierten Fahren: der Hersteller oder der Nutzer?
- Führt automatisiertes Fahren zur Reduktion des Verantwortungsbewusstseins und zu fahrlässigem Handeln der Nutzer?
- Wollen wir eine Fremdsteuerung durch Technologie?
- Führt Smart Mobility zu Gängelung oder zu Ermächtigung?
- Ist Autofahren eine private Angelegenheit, und sind die Daten, die dabei entstehen, privat?
- Sollten die Daten beim Autofahren automatisch übermittelt und gespeichert werden, oder sollte der Nutzer eine Wahlmöglichkeit haben?
- Sind Geschäftsmodelle mit Versicherungstarifen, die auf Grundlage von Daten angepasst werden, im Bereich Mobilität sinnvoll?
- Welche Rolle spielt Freiheit beim Autofahren?
- Ist die Sicherheit im Verkehr wichtiger als die individuelle Freiheit beim Autofahren?
- Kann Big Data im Mobilitätsbereich den Verkehr effizienter machen und die Sicherheit erhöhen?

Smart Health

- Erhöhen Produkte wie Wearables die Ich-Bezogenheit und Individualität in der Gesellschaft?
- Leben Nutzer von Produkten wie Wearables wirklich gesünder?
- Müssen Produkte wie Wearables die Möglichkeit bieten, dass die Datenübermittlung an den Anbieter ausgeschaltet werden kann?
- Profitiert die Gesellschaft von einer Individualisierung von Gesundheitsprodukten wie etwa Krankenkassentarifen?
- Sollte das Solidaritätsprinzip im Gesundheitssystem überdacht werden?
- Kann Smart Health das Gesundheitssystem ökonomisch effektiver machen?
- Ist eine Datenerhebung zur besseren Einschätzung von beispielsweise komplizierten Krankheitsbildern sinnvoll?

- Können durch Smart Health Kosten im Gesundheitssystem gespart werden?
- Ist eine Datenerhebung zur besseren Einschätzung von beispielsweise komplizierten Krankheitsbildern sinnvoll?
- Sollten Patienten selbst entscheiden können, ob ihre Daten zu Forschungszwecken freigegeben werden?
- Ist die Technik überhaupt so weit, dass die Daten einen Fortschritt für die Medizin bedeuten würden?
- Sollten bestimmte Daten im Gesundheitsbereich grundsätzlich nicht erhoben und gespeichert werden?
- Sind Gesundheitsdaten private Daten?
- Inwieweit kommen die Daten von Smart-Health-Produkten bisher überhaupt zum Einsatz? Werden sie von Krankenkassen genutzt oder nicht?
- Brauchen wir mehr staatliche Regulierung in diesem Bereich?

Die Expertenrunde war sich darin einig, dass es zur Beantwortung der aufgeworfenen Fragen einer breiten, gesamtgesellschaftlichen Debatte bedürfe. Die Diskussion konnte insoweit einen ersten Beitrag leisten, als die Experten konkrete Fragen und zum Teil Lösungsvorschläge formuliert haben, die als Grundlage für den weiteren Projektverlauf sehr wertvoll waren.

Die erste öffentliche Veranstaltung im Projektverlauf sollte die Debatte um Big Data in einen gesamtgesellschaftlichen Diskurs einbetten. Am 25. Juni 2015 diskutierten geladene und interessierte Gäste in der Stuttgarter Phönixhalle zur Frage „Vernetzte Mobilität – Erweiterung der persönlichen Autonomie oder Eingrenzung der Privatsphäre?“. In drei Keynotes und einer abschließenden Podiumsdiskussion erörterten Experten aus Politik, Wirtschaft und Wissenschaft den Einsatz von Big Data im Bereich Mobilität. Im Speziellen ging es um die Chancen und Risiken von selbst fahrenden Autos, Verkehrsleitsystemen sowie der Sammlung und Nutzung der dabei anfallenden Daten. Die Veranstaltung zeigte, dass die meisten Akteure davon ausgehen, die Vernetzung im Bereich der Mobilität werde wirtschaftlich und gesellschaftlich positive Effekte zeitigen. Es müsse

allerdings sichergestellt sein, dass der Mensch stets im Zentrum der Entwicklung verbleibe. Er müsse in die Lage versetzt werden, das Für und Wider von Big-Data-Anwendungen abzuwägen und darauf aufbauend selbstbestimmte Entscheidungen zu treffen. Zudem herrschte in Stuttgart weitgehend Einigkeit, dass bisher bestehende rechtliche Regelungen wie zum Beispiel das Datenschutzrecht für viele Fragen im Zusammenhang mit Big Data nicht zeitgemäß und hinreichend erscheinen.

„Das Datenschutzrecht wird oft als Universalwaffe genutzt – auch dann, wenn die Dinge gar nichts mit Datenschutz zu tun haben.“

Prof. Dr. Nikolaus Forgó,
Professor für IT-Recht und
Rechtinformatik an der Leibniz
Universität Hannover, öffentliche
Veranstaltung Stuttgart, 25.06.2015

Am 30. September fand die zweite öffentliche Veranstaltung des Teilprojekts statt. Im Meistersaal in Berlin befassten sich die eingeladenen Experten aus Wirtschaft, Medizin, Politik und Wissenschaft unter dem Titel „Treuer Assistent oder Trojaner am Körper? –

Wie Gesundheitstracking unseren Alltag verändert“ mit dem Thema der zunehmenden Datennutzung und Vernetzung im Gesundheitswesen. Ein besonderer Schwerpunkt lag dabei auf dem Thema „Tracking mit Wearables“ und den damit verbundenen Chancen und Risiken. Dabei zeigte sich während der Keynotes und der Diskussionen zwischen den Teilnehmern, dass die verstärkte Nutzung von Gesundheitsdaten, die mit der Einführung der neuen Technologien einhergeht, große Potenziale bietet. Beachtliche medizinische Fortschritte seien zu erwarten, zudem könnte die Digitalisierung in der Medizin eine bessere Versorgung des einzelnen Patienten und ein effektiveres Gesundheitssystem zeitigen. Zugleich aber dürften die Risiken nicht aus den Augen verloren werden. Besonders die zunehmende private Nutzung von Wearables und anderen Methoden zum Tracking der eigenen Gesundheitsdaten berge die Gefahr, dass sich die Leistungsgesellschaft noch verschärfe, was letztlich eine Entsolidarisierung zur Folge haben könnte. Zumindest aber sei eine grundsätzliche gesellschaftliche Debatte zu dem Thema unbedingt notwendig.

Neben den Diskussionsrunden und den öffentlichen Veranstaltungen wurde das Projekt von einer Reihe von Interviews und Konsultationen über Big Data, Smart Health und Smart Mobility begleitet, die mit Akteuren und Stakeholdern geführt wurden. Die Interviews sind Teil dieses Berichts.

ZEITSTRAHL ZUM PROJEKTVERLAUF „BIG DATA“

SEPTEMBER 2014

Projektbeginn

FEBRUAR 2015

Auftaktworkshop der
Expertengruppe in Berlin

Veröffentlichung des
Themenpapiers „Big Data“

JUNI 2015

2. Expertenworkshop zu Smart Health
und Smart Mobility unter Beteiligung
von Akteuren in Berlin

Öffentliche Diskussionsveranstaltung
in Stuttgart: „Vernetzte Mobilität –
Erweiterung der persönlichen Autonomie
oder Eingrenzung der Privatsphäre?“

JULI 2015

Veröffentlichung des Themenpapiers
„Gesellschaftliche Konfliktlinien
im Kontext von Big Data
am Beispiel von Smart Health
und Smart Mobility“

SEPTEMBER 2015

Öffentliche Diskussionsveranstaltung
in Berlin: „Wearables & Big Data“

JANUAR 2016

Veröffentlichung des Projektberichts

1.7 Aufbau des Dokuments

Der vorliegende Bericht ist entlang der Linien aufgebaut, wie sie durch die beiden Teilbereiche Smart Mobility und Smart Health vorgegeben werden. Nach der einleitenden Geschichte folgt jeweils die Definition des Unterthemas, das anschließend anhand von insgesamt drei Beispielen – telematikbasierte Autoversicherung und Verkehrslenkung bei Smart Mobility, Wearables wie Fitnessarmbänder oder Smart Watches bei Smart Health – näher dargestellt wird

mit einem Fokus auf den Chancen und Risiken des Anwendungsfalles. Die Kapitel schließen jeweils mit einer Darstellung der Herausforderungen, die sich daraus für die Gesellschaft ergeben. Das vierte Kapitel bietet ausgehend von den drei Beispielen einen Ausblick auf mögliche Kodifizierungen. Es werden die übergreifenden Aspekte herausgearbeitet und die sich daraus ergebenden gesellschaftlichen Grundkonflikte benannt. Der Bericht endet mit einem Ausblick auf einen möglichen weiteren Projektverlauf.

2 Smart Mobility: Tracking im Bereich Mobilität

Stellen Sie sich folgendes Szenario vor, das in naher Zukunft schon Wirklichkeit sein könnte:

Sie fühlen sich großartig! Zum ersten Mal seit mehreren Monaten sind Sie ohne Stau von einem Termin in der Firmenzentrale nach Hause gekommen. Das neue Big-Data-Dashboard auf dem Navigationsbildschirm Ihres Autos zeigt Ihnen nicht nur Verkehrsdaten, Stauwarnungen, Wetterdaten, Tankstellenpreise, Baustellen und Alternativrouten. Es verrechnet sogar die Benutzerdaten aller Nutzer in Echtzeit und verhindert durch entsprechend koordinierte Empfehlungen an die Verkehrsteilnehmer, dass ein Stau erst entsteht. Sie sind vollauf zufrieden, dass Sie der neuen App den Zugang auf die Bewegungsdaten Ihres Autos gewährt haben. Mobile Radarkontrollen werden von anderen Nutzern der App gekennzeichnet, die Öffnungszeiten von Rasthöfen werden direkt angezeigt, und bei Regen werden Ihnen Straßen vorgeschlagen, bei denen es weniger wahrscheinlich zu Aquaplaning kommt. Dank der neuen Software ist das Autofahren für Sie wesentlich komfortabler und sicherer geworden.

Im Rahmen eines öffentlichen Forschungsprojektes melden Sie sich mit einer ähnlichen App für eine Untersuchung an, bei der Daten gesammelt werden, die von verschiedenen Behörden zur Stadtplanung und Verkehrsregulierung benutzt werden. Von einem ersten Ergebnis profitieren Sie direkt: Die Ampeln in Ihrer Nachbarschaft reagieren in Echtzeit auf die zahlreichen Teilnehmer der Untersuchung, wodurch die Wartezeiten erheblich verkürzt werden. In der da-

zugehörigen Software können Sie alle Effekte nachverfolgen. So hat sich in Ihrer Stadt der CO₂-Ausstoß seit Beginn der Untersuchung merklich verringert, weil weniger Autofahrer im Stau stehen und die Gesamtdauer der kollektiven Autonutzung zurückgeht.

Es gibt auch deutlich weniger Unfälle, weil viele Teilnehmer bewusster fahren. Sie freuen sich über die höhere Sicherheit im Straßenverkehr und lassen ihre Kinder seitdem ruhigen Gewissens alleine zum Kiosk um die Ecke gehen. Die Anzahl der Autodiebstähle ist zurückgegangen und die Aufklärungsquote dennoch begangener Diebstähle ist dank der Ortungsmöglichkeit der Fahrzeuge sehr hoch.

Aufgrund der positiven Erfahrungen mit dem Tracking Ihres Autos melden Sie sich für eine telematikbasierte Kfz-Versicherung an, bei einem „Pay how you drive“-Programm. Dabei werden Ihre Versicherungsbeiträge nicht länger ausschließlich auf der Basis traditioneller Faktoren wie Alter, Geschlecht, Wohnort oder bisheriger Versicherungsfälle berechnet. Die Höhe der neuen Versicherungspolice wird auch von Ihrem Fahrverhalten beeinflusst: Geschwindigkeit, zurückgelegte Entfernung pro Fahrt, Nachtfahrten, häufiges starkes Bremsen oder ausgeglichenes, umsichtiges Dahingleiten. All diese Dinge wirken sich direkt auf Ihre monatlichen Abbuchungen aus. Der Vorteil ist besonders groß für Ihren Sohn, der als Fahranfänger eine recht teure Versicherung hat. Durch vorsichtiges Fahren kann er die Kosten immerhin ein Stück weit reduzieren. Als positive Nebenwirkung Ihrer eigenen ruhigeren Fahrweise verringern

sich die Kosten für Verschleißteile wie Reifen und Bremsen. Schade ist nur, dass sich der Tarif für Ihre Schwester nicht lohnt. Weil sie im Schichtdienst arbeitet, ist sie häufig nachts unterwegs, was sich unvorteilhaft auf die Preisbestimmung auswirkt.

Damit die Versicherungsgesellschaft auf Ihre Fahrdaten zugreifen kann, haben Sie eine sogenannte Blackbox in Ihrem Auto installiert, die alle Daten misst und übermittelt. Nach ein paar Wochen mit dem neuen Spielzeug beschwert sich Ihr Sohn. Er findet es stressig, für jede etwas größere Beschleunigung direkt die Strafe in Form einer Preiserhöhung zu bekommen. Das Freiheitsgefühl beim Autofahren ist auch Ihnen abhandengekommen. Sie fühlen sich in Ihrem Auto nicht mehr wie der Kapitän an Bord, sondern wie jemand, der stets darauf achten muss, die Gebote der Blackbox zu beachten.

Ein Bekannter hat Ihnen erzählt, dass die gleiche Versicherung bei Firmenwagen ein neues Konzept anbietet. Demnach würde die Box bei längeren Fahrten auf Pausen hinweisen und bei Missachtung der Hinweise beispielsweise das Radio und die Klimaanlage deaktivieren, um den Fahrer zu seiner „empfohlenen“ Pause zu bewegen. Die Pausenintervalle könnte dabei der Arbeitgeber festlegen. Über die Benutzung der Firmenwagen wäre der Arbeitgeber ohnehin informiert. Bewegungsprofile der Angestellten würden zwar nicht an den Arbeitgeber übermittelt, aber man wisse ja, wie es mit der Einhaltung von Regeln sei, sobald die Daten erst einmal vorhanden seien, meint ihr Bekannter. Er ist der Meinung, wenn sich erst alle an die Benutzung und an die Vorteile der Blackbox gewöhnt hätten, würde keiner das Ding wieder hergeben, nur weil es uns ein bisschen hinterherspioniert!

Weil Ihr Bekannter für eine Behörde arbeitet, kennt er einige Überlegungen für eine öffentliche Nutzung solcher Blackboxes. Er erzählt Ihnen von der Idee, die Höchstgeschwindigkeit von Autos über die Box zu deckeln, um eine höhere Verkehrssicherheit zu erreichen. Die App könnte außerdem die SMS- und Messaging-Funktionen des Fahrer-Smartphones automatisch blockieren, sobald das Auto in Bewegung sei. Bei diesen Gedanken fühlen Sie sich zunächst ziemlich bevormundet. Wenn Sie aber an Ihren Sohn denken, gefällt Ihnen die Idee dann doch ganz gut.

Eines Morgens erscheint eine Nachricht in Ihrer Versicherungs-App. Es ist eine Warnung. Ihr Sohn

hatte am vergangenen Wochenende offenbar das Auto von 22 Uhr bis 6 Uhr morgens vor einer Diskothek geparkt. Danach wurde das Auto kurz bewegt und kaum drei Stunden später mit seinem individualisierten Autoschlüssel schon wieder gefahren. Die App weist darauf hin, dass Fahrten in übermüdetem Zustand ein Sicherheitsrisiko darstellen. Sie wundern sich über diese genauen Informationen und stellen fest, dass Sie die gesamten Bewegungsprofile aufgeschlüsselt nach dem jeweiligen Fahrer einsehen können.

Sie ertappen sich dabei, sich überwacht zu fühlen. Wenn Sie abends spät noch Einkäufe erledigen müssen, vermeiden Sie den nächstgelegenen Supermarkt, weil allgemein bekannt ist, dass dort auf dem Parkplatz Drogen verkauft werden. Sie fahren lieber ein paar Kilometer weiter, um keine Bewegungsdaten zu erzeugen, die irgendjemand falsch interpretieren könnte. Schließlich sind Sie nicht sicher, wer alles Zugriff auf Ihre Daten haben könnte.

Ein paar Tage später lesen Sie in der Zeitung, dass es einen neuen Gesetzesvorschlag gebe, nach dem die Unschuldsvermutung für Verkehrsdelikte nur noch dann gelten soll, wenn die Verkehrsteilnehmer über ihre Blackbox oder im Falle von Radfahrern über Tracking-Armbänder nachweisen könnten, dass sie sich korrekt verhalten haben. Ohne eine solche Tracking-Technik könnte nach einem Unfall ja jeder behaupten, er hätte gebremst! Wer sich zum Beispiel als Radfahrer einer Analyse der gyroskopischen Aufzeichnungen seines Armbands entziehen wolle, sei ohnehin verdächtig, schreibt die Arbeitsgruppe, die den Gesetzentwurf vorbereitet hat. Die Weigerung, Tracking-Technologie zu nutzen, könnte laut dem Vorschlag in Zukunft als Gefahr für die öffentliche Sicherheit gelten. Kleinere Vergehen, wie Überschreitungen der Höchstgeschwindigkeit, könnten dank Tracking direkt registriert werden, und das entsprechende Bußgeld würde automatisch auf die Steuerzahlungen aufgeschlagen werden.

Zusätzlich gibt es Pläne, die Daten aus der öffentlichen Untersuchung zum Stadtverkehr, an der Sie teilnehmen, zu verwenden, um Verhaltens- und Charakterprofile der Verkehrsteilnehmer zu erstellen, die nach festgelegten Kriterien in Gerichtsverfahren einfließen könnten. Das gefällt Ihnen überhaupt nicht. Damit sind Sie nicht einverstanden. Als Sie jedoch die Teilnahmebedingungen für die öffentliche Untersu-

chung genauer lesen, stellen Sie fest, dass Sie sich damit schon einverstanden erklärt haben.

Einige Zeit später sind diese Vorschläge umgesetzt. Weil Sie eine rote Ampel überfahren haben, bekommen Sie von der öffentlichen Verkehrssoftware einen zusätzlichen Punkt in Flensburg. Leider tritt in Ihrem Fall ein Fehler auf, der auch bei einigen Hundert anderen Führerscheininhabern auftritt. Im Zuge der automatischen Punktevergabe wird Ihr persönliches Fahrerprofil fälschlicherweise als gefährlich eingestuft, und es wird Ihnen die Fahrerlaubnis entzogen. Ihr eigenes Auto reagiert nun nicht mehr auf Ihren Schlüssel, und weil das Lenkrad ähnlich wie ein Fitnessarmband Ihre Körperdaten analysiert, können Sie auch nicht einfach mit dem Autoschlüssel Ihres Sohnes fahren. Sie haben also keine andere Wahl, als den Fehler der zuständigen Behörde zu melden. Weil diese den Algorithmus für die Erstellung der Fahrerprofile jedoch bei einer privaten Firma eingekauft hat, ist sie nicht unmittelbar handlungsfähig. Die private Firma befindet sich gerade in einem kartellrechtlichen Streit über genau den Algorithmus, der Ihnen zu Unrecht den Führerschein entzogen hat. Die Entwickler haben zudem das Unternehmen gewechselt, und das aktuelle Team hat Schwierigkeiten, den Fehler zu finden. Sie wissen nicht, wie lange die Bearbeitung Ihres Problems in Anspruch nehmen wird.

Aber so lange zeigt Ihnen die App auf Ihrem Handy immerhin die Verfügbarkeit und die Preise der nächstgelegenen öffentlichen Verkehrsmittel und Taxis an.

2.1 Begriffsdefinition

Das beschriebene Szenario ist bislang noch Zukunftsmusik. Das heißt jedoch nicht, dass es sich bei der Geschichte um reine Science-Fiction handeln würde, die völlig aus der Luft gegriffen wäre. Im Gegenteil: Viele der genannten Technologien existieren bereits, befinden sich in der Testphase oder wurden von Forschern und Ingenieuren zumindest konzipiert. Sie alle

lassen sich unter dem Begriff Smart Mobility zusammenfassen, um den es in den folgenden Abschnitten gehen soll.

„Ich sehe großen Nutzen in der neuen, digitalen Mobilität: Gefahren verhindern, Verkehrseffizienz verbessern und Komfort erhöhen.“

Parlamentarischer Staatssekretär beim Bundesminister für Verkehr und digitale Infrastruktur,
Norbert Barthle, MdB, öffentliche Veranstaltung Stuttgart, 25.06.2015

Smart Mobility wird als die Zusammenfassung all jener Technologien umschrieben, die zukunftsorientiert darauf ausgerichtet sind, den Bürgern einer Stadt, einer Region oder eines Landes Mobilität zu ermöglichen, die zugleich effizient, umweltschonend, emissionsarm, komfortabel, sicher und kostengünstig ist. Diese Ziele sollen insbesondere durch den Einsatz von Informations- und Kommunikationstechnologien erreicht werden.¹⁸ Mobilität ist dabei umfassend zu verstehen und nicht auf ein bestimmtes Verkehrsmittel beschränkt. Vielmehr geht es grundsätzlich gerade darum, alle verfügbaren Mobilitätsformen zu vernetzen, um die Mobilität aller Bürger zu verbessern und zu optimieren. Das Phänomen als solches ist räumlich nicht auf die Stadt beschränkt. Die meisten der verfügbaren Anwendungsbeispiele beziehen sich jedoch auf den urbanen Raum als primärem Planungsgegenstand politischer Strategien. In diesem Sinne wird das Themenfeld neben Smart Governance/Smart Education, Smart Health, Smart Building, Smart Infrastructure, Smart Technology, Smart Energy und Smart Citizen als einer der acht Grundpfeiler, des übergeordneten Handlungsziels Smart City identifiziert.¹⁹ Unter Smart City versteht man den Einsatz von Informations- und Kommunikationstechnologien in urbanen Konglomerationen, um die vorhandenen Ressourcen intelligenter und effizienter zu

¹⁸ Vgl. Stefan Wolter, „Smart Mobility – Intelligente Vernetzung der Verkehrsangebote in Großstädten“, Vortrag an der Universität Düsseldorf, Abstract online: https://www.uni-due.de/imperia/md/content/iam/track2_abstract_wolter.pdf.

¹⁹ Sarwant Singh, „Smart Cities – A \$1.5 Trillion Market Opportunity“, Forbes Online, 19. Juni 2014, online: <http://www.forbes.com/sites/sarwantsingh/2014/06/19/smart-cities-a-1-5-trillion-market-opportunity/>.

nutzen. Daraus sollen Kosten- und Energieeinsparungen und zugleich eine Erhöhung der allgemeinen Lebensqualität resultieren.²⁰ Die Fokussierung auf Städte hat insofern Sinn, als allgemein vorausgesagt wird, dass bis zum Jahr 2050 drei Viertel der Weltbevölkerung in Städten leben werden.²¹

Der Begriff Smart Mobility soll in diesem Dokument allerdings weiter gefasst werden und auch datengetriebene Anwendungen in Mobilität unterstützenden Wirtschaftszweigen beinhalten, beispielsweise telematikbasierte Versicherungstarife. Auch in diesem weiter gefassten Begriffsverständnis kommt der Nutzung von Big Data für die Umsetzung der Ziele von Smart Mobility eine Schlüsselrolle zu. Erst durch die Sammlung, Zusammenführung und Auswertung von Daten aus verschiedenen mobilitätsbezogenen Quellen wird Mobilität „intelligent“. Durch die Analyse strukturierter und unstrukturierter Daten können die beteiligten Akteure Strategien entwickeln, um die Ressourcennutzung zu optimieren.

Traditionelle Datenquellen im Bereich Mobilität sind zeit- und ortsabhängige Daten wie Fahr- und Zeitpläne des öffentlichen Personennahverkehrs, Stadtkarten, Informationen zu Veranstaltungen, Unfällen oder Staumeldungen, Daten wie Statistiken zur Auslastung von Straßen und Verkehrsmitteln, Preise für öffentlichen Personennahverkehr oder Taxis, Feedback von Kunden öffentlicher Verkehrsmittel und schließlich sogar Informationen bezüglich für sich genommen mobilitätsferner Umstände wie Berichte über politische Ereignisse oder Wettervorhersagen.²² Durch moderne Informations- und Kommunikationstechnologien in jüngerer Zeit hinzugekommen sind außerdem Datenquellen, die mobilitätsrelevante Vorkommnisse in Echtzeit erfassen und deren IT-gesteuerte Auswertung eine zeitnahe oder sogar gleichzeiti-

ge Reaktion ermöglicht, um das Verkehrsgeschehen ressourcenoptimiert zu halten. Solche Quellen sind beispielsweise die aus Sozialen Medien gewonnenen unstrukturierten Daten wie Twitter-Meldungen von Betroffenen über aktuelle Staus oder Verspätungen im öffentlichen Nahverkehr.²³ Eine spezielle Anwendung für solche Zwecke ist die Smartphone-App Waze, ein Navigationssystem, das ein eigenes Soziales Netzwerk enthält, wodurch gezielt verkehrsrelevante Daten ausgetauscht und ins System eingespeist werden können. Das Unternehmen, das inzwischen von Google aufgekauft wurde, ist auch Kooperationen mit mehreren Städten eingegangen.²⁴ Wenn es gelingt, all diese Daten kontinuierlich und simultan zu analysieren, dann können im Idealfall kritische Situationen von vornherein verhindert werden. Die Computerlinguistin Feiyu Xu vom Deutschen Forschungszentrum für Künstliche Intelligenz nennt in ihrem beim Bitkom Big Data Summit 2015 gehaltenen Vortrag „Smart Data for Mobility“ folgendes theoretische Anwendungsszenario als Beispiel für einen Fall einer mobilitätsbezogenen Big-Data-Analyse in Echtzeit:

„Vor dem Frankfurter Bahnhof ist Wochenmarkt, und demnächst geht ein Fußballspiel in der Umgebung zu Ende. Weiterhin zieht ein Unwetter auf, und die Menschen suchen Schutz im Bahnhof. Durch Auswertung von Wetter- und Social-Media-Daten wurden im Vorfeld mehr Sicherheitskräfte aus den umliegenden S-Bahnen in den Frankfurter Hauptbahnhof disponiert.“²⁵

Dieses Beispiel zeigt anschaulich, wie aus verschiedenen Datenquellen (Veranstaltungskalender, Wettervorhersage, Social Media) Informationen zusammengeführt werden, um auf aktuelle Ereignisse schnell und in angemessener Weise reagieren zu können.

²⁰ Boyd Cohen, „The Top 10 Smart Cities on the Planet“, Fast Company, 11. Januar 2012, online: <http://www.fastcoexist.com/1679127/the-top-10-smart-cities-on-the-planet>.

²¹ Jane Wakefield, „Tomorrow's cities: Do you want to live in a smart city?“, BBC Online, 19. August 2013, online: <http://www.bbc.com/news/technology-22538561>.

²² Vgl. Feiyu Xu, „Smart Data for Mobility (SD4M). Eine Big-Data-Analytik-Plattform für multimodale Smart Mobility Services“, Vortrag beim Bitkom Big Data Summit 2015, 25. Februar 2015, online: http://www.bitkom-bigdata.de/sites/default/files/14.30%20BDS15%20Schwarzer%20DB%20Systel_DrXU%20DFKI_0.pdf.

²³ Ebd.

²⁴ Vgl. Ryan Bradley, „Waze and the Traffic Panopticon“, The New Yorker, 2. Juni 2015, online: <http://www.newyorker.com/business/currency/waze-and-the-traffic-panopticon>.

²⁵ Xu, „Smart Data for Mobility“, s.o.

INTERVIEW MIT MARIT HANSEN

Es besteht eine Lücke zwischen dem Datenschutz und der Technik

? **Big Data ist der Taktgeber der digitalen Revolution. Was hat sich im Verhältnis zum Umgang mit Daten im analogen Zeitalter eigentlich verändert?**

Marit Hansen: Bei Big Data wird aus einer quantitativen Veränderung, nämlich dass vermehrt große Mengen an Daten vorhanden sind und ausgewertet werden können, eine qualitative. Diese Daten kommen in unterschiedlichen Formaten und aus unter-

schiedlichen Quellen. Aus Datenschutzsicht ist dies bedeutsam, weil sich damit die Zweckbindung, die bisher im Datenschutzrecht eine große Rolle gespielt hat, nicht mehr durchhalten lässt. Bisher waren Datenbanken begrenzt: durch ihre For-

Marit Hansen

Marit Hansen ist seit 2015 die Landesbeauftragte für Datenschutz Schleswig-Holstein und leitet das Unabhängige Landeszentrum für Datenschutz (ULD). Davor war die Diplom-Informatikerin sieben Jahre lang stellvertretende Landesbeauftragte für Datenschutz. Im ULD hat sie den Bereich der Projekte für technischen Datenschutz aufgebaut.

Seit 1995 arbeitet sie zu Themen des Datenschutzes und der Informationssicherheit. Ihr Schwerpunkt liegt auf der grundrechtskonformen Gestaltung von Systemen, insbesondere durch Privacy by Design und Privacy by Default. Privacy- und Datenschutzfragen hat sie in zahlreichen Publikationen und Vorträgen behandelt.

Foto: Markus Hansen



“ Wenn ich das geltende Recht befolgen möchte, kann ich viele der Big-Data-Anwendungen nicht einsetzen.

Marit Hansen

mate, durch den Kontext, durch den Zweck der Sammlung. Heute werden diese Daten mit anderen zusammengelegt und gemeinsam interpretiert.

? **Bedeutet das, es werden Daten gesammelt, ohne dass man weiß, was man damit tun wird?**

MH: Genau. Man weiß noch gar nicht, welche Informationen in den Daten enthalten sind. Man gibt alles in einen Computer und lässt ihn vor sich hin rechnen. Heraus kommen dann Korrelationen, die nutzbar sind. Zum Beispiel: Wie organisiert man am besten Waren im Supermarkt, so dass man die Kunden dazu bringt, möglichst viel einzukaufen? Die Erwartung ist, dass man durch die Big-Data-Analyse Erkenntnisse gewinnt, die vorher nicht absehbar waren. Deswegen kann das Ziel der Verwendung nicht

genau festgelegt werden. Ausgangspunkt sind eher allgemeine Fragestellungen, zum Beispiel: Wie kann ich Geld einsparen? Wie kann ich Ressourcen besser verteilen? Konkret etwa: Wie verteile ich medizinisches Personal möglichst effektiv? Oder im Bereich Kriminalitätsprävention: Wer wird in dieser Zeit an welchem Ort eine kriminelle Tat begehen und wie kann ich das verhindern?

? **Sind Ihrer Meinung nach die rechtlichen Regelungen für Big Data ausreichend? Oder brauchen wir neue Gesetze?**

MH: Wenn ich das geltende Recht befolgen möchte, kann ich viele der Big-Data-Anwendungen nicht einsetzen. Wenn wir bisher etwas über die Bevölkerung wissen wollten, hatten wir die Bundes- und Landesstatistikämter. Sie können verbindlich Zahlen herausfinden

über Dinge, die für den Staat und die Gesellschaft wichtig sind. Sie sind aber strengen Regeln unterworfen. Wer im Statistikamt arbeitet, ist zur besonderen Geheimhaltung verpflichtet. Ein Verstoß gegen das Statistikgeheimnis kann dazu führen, dass man ins Gefängnis kommt – es stehen hohe Strafen darauf.

Bei Big Data werden diese Regelungen nicht zum Einsatz gebracht – ich kenne keinen Fall. Es wird meist argumentiert: Die Daten sind bereits da und werden nicht neu erhoben. Dieses Vorgehen ist aber meist nicht rechtlich einwandfrei. Wenn die Daten nicht personenbezogen sind, könnte man darüber reden, inwieweit die Sammlung und Zusammenführung der Daten in Ordnung ist. Aber die meisten Big-Data-Anwendungen dürften gar nicht so weit kommen, weil die betreffenden Daten doch einen Personenbezug beinhalten.



ES BESTEHT EINE LÜCKE ZWISCHEN DEM DATENSCHUTZ UND DER TECHNIK



Die Frage, die sich dabei stellt, ist: Will man solche Anwendungen ermöglichen? Berauben wir uns sonst bestimmter Werkzeuge, die wichtig sind für unser Leben? Dann müssen wir uns überlegen, welche Vorkehrungen wir treffen müssen, damit bestimmte Anwendungen möglich sind.

? **Wie ist Ihre persönliche Bewertung? Brauchen wir ein Big-Data-Ermöglichungsgesetz oder sind diese Begrenzungen richtig?**

MH: Die Entwicklung geht dahin, dass Big-Data-Anwendungen einfach gemacht werden, ohne sich um die rechtlichen Einschränkungen zu kümmern. Eine Möglichkeit wäre, in technische Lösungen zu investieren, die den Personenbezug entfernen – als Informatikerin weiß ich, dass es sehr gute Ansätze gibt, – teilweise noch im For-

schungslabor, teilweise schon einsatzbereit. Die Anreize dafür sind aber zurzeit noch zu gering: Viele Anbieter machen einfach mit dem Argument weiter, dass das doch nützliche Anwendungen sind. Die Politiker und zuständigen Behörden setzen sich dagegen zu wenig zur Wehr.

Das bedeutet, dass wir uns als Gesellschaft überlegen müssen, was wir mit Big Data erreichen wollen und wie wir das Risiko beherrschen können. Diese Abwägung fehlt mir in der Diskussion oft noch. Auf der einen Seite können sich Datenschützer darauf zurückziehen, es sei sowieso alles illegal, also müsste man sich damit nicht weiter beschäftigen. Die Big-Data-Anbieter und -Anwender auf der anderen Seite sind anscheinend der Meinung: „Wir machen es in jedem Fall und müssen uns nicht damit beschäftigen, wie wir das Risiko ein bisschen minimieren.“

? **Um das an einem konkreten Beispiel zu diskutieren: Inwieweit darf man personenbezogene Daten aus den Sozialen Netzwerken nutzen? Wo sind für Sie die Grenzlinien?**

MH: Das ist eine schwierige Frage, weil es ganz verschiedene Social-Media-Dienste gibt, die nach unterschiedlichen Regeln funktionieren. Twitter funktioniert anders als Facebook. Bei Twitter stelle ich in der Regel etwas für die ganze Welt bereit, während ich bei Facebook auch in kleinen Gruppierungen agieren kann. Damit sind Erwartungen verknüpft, dass diese Daten und Inhalte auch in dieser Gruppe bleiben. Das ist nichts Neues: Wenn wir eine Konferenzschaltung machen, wäre das etwas anderes als ein Plakat im öffentlichen Raum oder eine Anzeige in der Zeitung.

Ein anderer Effekt, der oft nicht beachtet wird, ist, dass Soziale

“ Ein anderer Effekt, der oft nicht beachtet wird, ist, dass Soziale Medien nicht repräsentativ sind.

Marit Hansen

Medien nicht repräsentativ sind. Nur weil etwas in den Sozialen Medien nicht auftaucht, bedeutet das nicht, dass es nicht da ist. Ein Beispiel: Es wird oft gesagt, dass Jugendliche sich nicht an politischen Diskussionen beteiligen. Ich weiß aber von vielen, dass sie solche Themen im Netz absichtlich nicht diskutieren – sie sind aber trotzdem politisch interessiert. In den Medien hieß es aber, unsere Jugendlichen hätten kein Interesse an Politik. Das heißt, es wurden für eine Studie Daten aus Sozialen Medien genommen, die einfach zu sammeln waren, und daraus etwas abgeleitet, ohne zu überlegen, ob das überhaupt repräsentativ ist.

So etwas kann auch in ganz anderen Bereichen passieren, zum Beispiel bei der Bluetooth-Erfassung des Verkehrs. Das Verfahren wird immer öfter eingesetzt, um Verkehrsnetze zu verbessern. Wer kein Bluetooth-Gerät (z.B. ein Smartphone) hat oder sein Bluetooth ausschaltet, taucht in der

Datenmenge nicht auf. Daraus könnte dann folgen, dass eine Straße nicht ausgebaut wird, weil ja anscheinend nicht so viele Leute dort leben. Das könnte dazu führen, dass nicht nur in den armen Vierteln einer Stadt keine Straßen gebaut werden – das sind die ohne Smartphones –, sondern auch in Vierteln, in denen Menschen wohnen, die besonders datenschutz- oder sicherheitsbewusst agieren und ihre Bluetooth-Geräte ausschalten. Solch eine Auswertung auf Basis von einfach verfügbaren Daten führt nicht nur zu Fehlentscheidungen wegen der verzerrten Datenbasis, sondern auch zu einer Diskriminierung.

? **Ist Mobilität ein besonders gefährdeter Bereich? Die ersten Autoversicherungen haben angefangen, Begehrlichkeiten für Daten zu entwickeln. Derzeit stehen Fragen**

wie „Welche Daten kann ich generieren? Zu was kann ich die verwenden?“ auf der Tagesordnung.

MH: Mobilität ist nicht nur wegen Big Data ein besonderer Bereich, sondern weil immer größere Bereiche unseres Lebens durch Mobilität gekennzeichnet sind. Mobilität hat immer mit Standortdaten und Bewegungsdaten zu tun. Sie liefert Erkenntnisse über Netzwerke: Wer kennt wen? Wer hat mit wem zu tun? Wie gestaltet jemand sein Leben? Das sind alles zentrale Fragen für die informationelle Selbstbestimmung.

Durch die mobile Kommunikation wachsen Leben und Mobilität enger zusammen, denn um per Mobilfunk erreichbar zu sein, müssen bestimmte Informationen über den Standort bekannt sein: Die Mobilfunkinfrastruktur muss bei der heutigen technischen Gestaltung immer wissen, in welcher Funkzelle man eingebucht ist. Die-



ES BESTEHT EINE LÜCKE ZWISCHEN DEM DATENSCHUTZ UND DER TECHNIK



se Standortdaten werden dabei von den Anbietern nicht nur kurzfristig gespeichert, sondern länger und dann auch weiter ausgewertet. Viele der Probleme, die auf uns zukommen werden, werden mit Mobilität zu tun haben, sei es bei der Verkehrsdatenauswertung, sei es beim vernetzten oder selbst fahrenden Auto.

? **Haben Sie ein Lösungsmodell parat, wie man an diesen hochkomplexen Bereich herangehen könnte?**

MH: Gerade in unserem Telekommunikationssystem liegt noch vieles im Argen, was den Datenschutz angeht. Im Prinzip müsste man das gesamte Handynetze neu ausrollen – das ist aber sehr unrealistisch. Das Handynetze wurde mit dem Schwerpunkt auf der Verbesserung von Verfügbarkeit und Geschwindigkeit ausgebaut. Sicherheit und Datenvermeidung

spielten keine Rolle, und das ist kaum mehr umkehrbar. Genau daran hängt, dass wir für die Mobilitätsdaten ein besonderes Risiko für die informationelle Selbstbestimmung sehen.

Es gibt viele Dienste, bei denen Lokalisierungsdaten ganz praktisch sind, zum Beispiel Navigationsdienste. Aber auch da kann man schauen: Welche Dienste nutze ich zur Lokalisierung? Das muss nicht über Telekommunikationsdienste laufen, das kann auch über GPS auf dem eigenen Gerät passieren, d.h., die Ortungsdaten werden nicht weitergegeben. Man kann darauf achten, dass die Interpretation dieser Daten im Endgerät stattfindet statt zentral auf den Cloud-Servern externer Anbieter.

Es gibt weitere Techniken, wie Daten genutzt werden können, ohne dass es gleich problematisch ist, zum Beispiel über die Auslastung. Man kann zum Beispiel feststellen, ob gerade irgendwo ein Volksfest ist, weil

sich besonders viele Leute in eine Funkzelle eingebucht haben. Das ist eine statistische Größe, die üblicherweise kein großes Risiko bedeutet. Solche Techniken, Daten ohne Personenbezug zu sammeln, gibt es eine ganze Menge. Es ist technisch möglich, große Datenmengen zu sammeln und aus ihnen Strukturen und Korrelationen abzuleiten, ohne dass diese Daten auf einzelne Personen zurückzuführen sind. Es gibt bis jetzt zu wenige Anwendungen – da könnte man noch sehr viel mehr machen.

? **Was wünschen Sie sich aus Ihrer persönlichen Erfahrung heraus für den gegenwärtigen Stand der Diskussion? Mehr Aufklärung? Mehr diverse Meinungen?**

MH: Ich wünsche mir, dass zunächst die Sorgen und Probleme

Was bedeutet es, in dieser Welt zu leben, in der ständig Entscheidungen aufgrund von irgendwelchen Datenströmen getroffen werden?

Marit Hansen


me ernster genommen werden und nicht im Sinne von „Das ist gut für die Gesellschaft“ abgetan werden. Gerade in der Industrie gibt es diese Herangehensweise: „So sensibel sind die Daten doch gar nicht – ihr mit eurem Datenschutzrecht und eure Datensparsamkeit seid nicht mehr zeitgemäß.“ Eine solche Einstellung kann nicht sein. Wir sehen an zahlreichen missbräuchlichen Nutzungen, wie wichtig es ist, bereits im Vorfeld auf Datenvermeidung und Datensparsamkeit zu setzen. Das muss nicht nur Big Data, sondern kann auch ein Hack eines Seitensprungportals²⁶ sein.

Es besteht zurzeit eine große Lücke zwischen dem, an dem in der Privacy-Forschung gearbei-

tet wird, und dem, was gesetzlich im Datenschutz diskutiert wird. Viele Forschungsideen sind vielleicht noch nicht einsatzreif, aber bereits so weit entwickelt, dass man in einem gemeinsamen Diskurs vermutlich große Fortschritte machen könnte. Wenn wir diese Lücke nicht schließen, haben wir in jedem Fall unbefriedigende Lösungen.

Und natürlich brauchen wir einen gesellschaftlichen Diskurs mit denjenigen, die die Daten liefern, nämlich den Menschen. Wie viel wissen sie darüber, was mit ihren Daten geschieht? Was bedeutet es, in dieser Welt zu leben, in der ständig Entscheidungen aufgrund von irgendwelchen Datenströmen getroffen werden? Wie gehen sie damit um? Es gibt einige, die noch sehr wenig da-

rüber wissen, andere wissen gerade so viel, dass sie sich ohnmächtig fühlen gegenüber der Technik. Dieses Problem müssen wir unbedingt angehen, damit für alle klarer wird, was genau passiert und wie man es beeinflussen kann. Dazu gehören auch Regeln zu datenschutzfreundlichen Voreinstellungen in Apps und Plattformen. Das würde ich in diesem Zusammenhang gerne weiter diskutieren.

Letztendlich gibt es keine Patentlösung, die immer passt. Man muss ein bisschen genauer hingucken, und dann können wir tatsächlich Datenschutzrisiken minimieren oder sogar ausschließen, ohne dass die Anwendungen, die Informationen sinnvoll nutzen, auf der Strecke bleiben. 

²⁶ Anm. der Redaktion: Das Seitensprungportal Ashley Madison wurde im Sommer 2015 gehackt. Die persönlichen Daten der 36 Millionen Nutzer wurden daraufhin im Internet veröffentlicht.

2.2 Der Einsatz von Tracking für Smart-Mobility-Anwendungen

Viele Big-Data-Anwendungen im Bereich Smart Mobility bauen auf Tracking als zentraler Technologie zur Gewinnung von Daten auf. Insbesondere Standortdaten werden auf diese Weise aufgezeichnet, gespeichert und übermittelt. Solche Geolokationsdaten²⁷ von Fahrzeugen oder Personen tragen entscheidend zur Erweiterung des Bestandes verfügbarer Datenquellen bei und machen somit die neuesten intelligenten Mobilitätslösungen erst möglich. Solche über GPS oder Mobilfunkersfassungen generierte Daten können benutzt werden, um individuellen Verkehrsteilnehmern in Echtzeit für sie relevante Informationen zukommen zu lassen. So können die Positionsdaten eines Pendlers mit weiteren mobilitätsbezogenen Daten aus anderen Quellen kombiniert werden, um aktuelle, für seine Route bedeutende Vorkommnisse auf das Display seines Smartphones zu spielen, zum Beispiel die Empfehlung, aufgrund eines plötzlich erhöhten Passagieraufkommens ein Umsteigen – um im genannten Beispiel zu bleiben – im Frankfurter Hauptbahnhof zu vermeiden. Darüber hinaus können Geolokationsdaten für eine grundlegendere Analyse von Verkehrsströmen innerhalb eines urbanen Raumes genutzt werden, um daraus Erkenntnisse für eine optimierte Verkehrsplanung und -lenkung zu gewinnen. In einer italienischen Fallstudie wurden beispielsweise die GSM-Daten²⁸ von 25 Millionen Telefongesprächen von 350.000 Personen zwischen dem 15. Oktober und dem 9. November 2012 und die GPS-Daten von 1,5 Millionen Autofahrten zwischen Februar und März sowie Juli und August 2012 mit den GPS-Daten der öffentlichen Verkehrsmittel der süditalienischen Stadt

Cosenza zusammengeführt, um einige zuvor formulierte Fragen bezüglich der Verkehrsströme in der Gemeinde zu beantworten.²⁹ Die Nutzung von Tracking-Daten zur Optimierung der Verkehrslenkung wird später in diesem Dokument anhand der Anwendungsfälle Stockholm und Toronto näher beleuchtet.

„Big-Data-basierte Mobilität sollte weniger darauf abzielen, Menschen oder Güter schneller von A nach B zu befördern, als vorhandene Ressourcen effizienter zu nutzen, um beispielsweise Energie zu sparen und Abgase zu reduzieren.“

Prof. Dr. Claudia Eckert, Leiterin des Fraunhofer-Instituts für Angewandte und Integrierte Sicherheit (AISEC) und Professorin für IT-Sicherheit an der Technischen Universität München, 2. Expertenworkshop 15.06.2015

Aber nicht nur Geolokationsdaten spielen beim Einsatz von Big Data für Smart-Mobility-Lösungen eine Rolle. Tracking-Technologien im weiteren Sinne, analog zu den Gesundheits- und Fitnessdaten, die von Wearables aufgezeichnet werden, finden sich insbesondere in praktisch allen Autos neueren Datums. Bis zu 70 Geräte und Sensoren registrieren und speichern bereits heute in vielen modernen Fahrzeugen verschiedene Informationen wie die gefahrene Geschwindigkeit, Kilometerleistung, Bewegungsprofile, Tages- und Nachtfahrten, Fahrten in unfallträchtigen Ballungszentren, Fahrstil (z.B. häufiges abruptes Bremsen oder starkes Beschleunigen), Fahrerwechsel, Teileverschleiß, Unfälle und Pannen, Wartungshäufigkeit und -zeitpunkt sowie selbst vorgenommene Eingriffe am Fahrzeug.

²⁷ Geolokation bezeichnet die Identifizierung des Aufenthaltsortes eines beliebigen Objekts (Quelle: <http://en.wikipedia.org/wiki/Geolocation>).

²⁸ GSM steht für „Global System for Mobile Communications“, ein Standard für voll-digitale Mobilfunknetze, der hauptsächlich für Telefonie, aber auch für leitungsvermittelte und paketvermittelte Datenübertragung sowie Kurzmitteilungen genutzt wird (Quelle: http://de.wikipedia.org/wiki/Global_System_for_Mobile_Communications).

²⁹ Roberto Trasarti u.a., „Big data analytics for smart mobility: a case study“, Pisa 2014, online: <http://ceur-ws.org/Vol-1133/paper-57.pdf>.

Auf diese Weise erzeugen moderne Autos 20 Gigabytes an Daten pro Stunde.³⁰ Dank Vernetzung können diese Daten dann über Funk an Dritte wie den Fahrzeughersteller, Vertragswerkstätten, Verkehrsleitzentralen oder Versicherungen übertragen werden. Prognosen zufolge soll es im Jahr 2020 bereits 250 Millionen auf diese Weise vernetzte Fahrzeuge geben, somit jedes fünfte Auto weltweit.³¹ Die so generierten Daten können ebenfalls zu Big Data zusammengeführt und in verschiedener Weise nutzbar gemacht werden. Im nächsten Abschnitt werden Autoversicherungen als ein Anwendungsfall der Verwendung dieser Daten vorgestellt und analysiert.

„Fahrzeugdaten aus vernetzten Autos sollten nicht ohne ausdrückliche Zustimmung an Dritte weitergegeben werden. Aber selbst die Zustimmung, sein Fahrverhalten tracken zu lassen, wird eventuell leichtfertig von den Nutzern gegeben, weil diese sich des Wertes ihrer Daten nicht bewusst sind.“

Prof. Dr. Tobias O. Keber,
Professor für Medienrecht und Medienpolitik, Institut für Digitale Ethik (IDE) der Hochschule der Medien Stuttgart, Interview

³⁰ Errol S. van Engelen, „Big Data Analytics to Improve Smart Mobility in the 21st Century“, LinkedIn Pulse, 5. Januar 2015, online: <https://www.linkedin.com/pulse/big-data-analytics-smart-mobility-21st-century-errol-s-van-engelen>.

³¹ René Kohlenberg, „Diese Dinge weiß Ihr Auto schon über Sie“, Kölner Stadtanzeiger Online, 10. Februar 2015, online: <http://www.ksta.de/service/digitale-daten-diese-dinge-weiss-ihr-auto-schon-ueber-sie,16126596,29799146.html>.

Öffentliche Veranstaltung in Stuttgart

VERNETZTE MOBILITÄT – ERWEITERUNG DER PERSÖNLICHEN AUTONOMIE ODER EINGRENZUNG DER PRIVATSPHÄRE?



Am 25. Juni 2015 fand in der Stuttgarter Phoenixhalle die fünfte öffentliche Veranstaltung im DIVSI-Projekt „Braucht Deutschland einen Digitalen Kodex?“ statt. In drei Keynotes und einer Podiumsdiskussion diskutierten Experten aus Politik, Wirtschaft und Wissenschaft den Einsatz von Big Data im Bereich der Mobilität, insbesondere die Chancen und Risiken von selbst fahrenden Autos, Verkehrsleitsystemen sowie der Sammlung und Nutzung dabei anfallender Daten. Im Zentrum standen neben den technologischen Möglichkeiten und neuen politischen Handlungsfeldern vor allem die gesellschaftlichen Herausforderungen.

Der Minister für Verkehr und Infrastruktur Baden-Württemberg Winfried Hermann stellte in seiner Begrüßung heraus, dass

die Automatisierung der Mobilität nicht als reiner Selbstzweck verstanden werden sollte. Die damit verbundenen Funktionen und Ziele müssten klar definiert sein. Beim autonomen Fahren seien beispielsweise die Reduktion von Unfällen und Klimaschäden am Gemeinwohl orientierte Funktionen, so der Minister. In Bezug auf die anfallenden Daten betonte Hermann die Rolle der Politik, die dafür Sorge tragen müsse, dass die Daten geschützt und zu klar definierten Zwecken und nicht nur im Interesse einiger privatwirtschaftlicher Unternehmen genutzt werden.

In der ersten Keynote erläuterte der Parlamentarische Staatssekretär beim Bundesminister für Verkehr und digitale Infrastruktur Nobert Barthle zunächst die Vorteile einer Digitalisierung im Be-

reich der Mobilität. Einen direkten Nutzen sieht Barthle etwa in Verkehrsleitsystemen, Carsharing und Notfalleinrichtungen über das sogenannte eCall-System. Die Verminderung von Gefahrenpotenzialen, eine Erhöhung der Verkehrseffizienz und ein steigender Komfort nannte der Staatssekretär in diesem Zusammenhang als positive Effekte. Darüber hinaus eröffneten sich neue Wachstumsperspektiven für die deutsche Wirtschaft. Jeder siebte Arbeitsplatz in Deutschland hänge von der Automobilindustrie ab, weshalb es wichtig sei, dass Innovationschancen und die daraus entstehenden Wertschöpfungsketten in diesem Bereich genutzt würden, so Barthle. Das Bundesministerium für Verkehr und digitale Infrastruktur habe diese Möglichkeiten erkannt und das Investitionsvolumen massiv



Podiumsdiskussion: (v.l.n.r.) Jörg Klingbeil, Landesbeauftragter für den Datenschutz Baden-Württemberg; Alexander Mankowsky, Zukunftsforscher Daimler AG; Lena-Sophie Müller, Geschäftsführerin der Initiative D21 e.V.; Ivo Körner, Geschäftsführer Vertrieb Branchenkunden IBM Deutschland GmbH; Ulrich Chiellino, Leiter Interessenvertretung Verkehr ADAC e.V.

erhöht, um eine internationale Konkurrenzfähigkeit zu erreichen. Zentral bei der Förderung durch das Ministerium sei dabei eine grundsätzliche Technologieoffenheit.

Alexander Mankowsky, Zukunftsforscher der Daimler AG, bot in seiner Keynote einen Einblick in die Arbeit der Entwicklung technischer Innovationen. Dabei betonte er, dass es zentral sei, sich nicht nur mit der Technologie zu beschäftigen, sondern zunächst vielmehr mit den gesellschaftlichen Auswirkungen neuer Vernetzungsmöglichkeiten. Der Einzug der Digitalisierung in den Mobilitätsbereich könne zu einer neuen Konsumgüterkultur führen, die das Potenzial hätte, die Gesellschaft zu verändern. Wie dann diese Gesellschaft aussehen würde, sei noch offen, wes-

halb sich die Zukunftsforschung bei Daimler etwa auch durch die Anwendung künstlerischer Methoden möglichen Szenarien annäherte. Eine Automatisierung und Vernetzung sei in vielen Bereichen

technisch zwar sehr umfassend möglich, zuvor aber sollte es immer eine Debatte darum geben, ob dies gesellschaftlich erstrebenswert wäre. Es müssten Wege gefunden werden, wie die techni- ➤



Minister für Verkehr und Infrastruktur Baden-Württemberg,
Winfried Hermann



Parlamentarischer Staatssekretär beim Bundesminister für Verkehr und digitale Infrastruktur, **Norbert Barthle**, MdB

› schen Automaten in die soziale Welt des Menschen auf sinnvolle Weise integriert werden könnten. Vertrauen zu schaffen sei dafür allein nicht ausreichend. Vielmehr müssten Prozessansätze wie Fairness, Transparenz und Kommunikation angewendet werden, so der Zukunftsforscher.

„If the product is for free, you are the product.“ Mit diesen Worten beschrieb Prof. Dr. Nikolaus Forgó vom Institut für Rechtsinformatik der Leibniz Universität Hannover in seiner Keynote die neue Welt der Gratisdienste und den damit einhergehenden Herrschaftsverlust. In diesem Zusammenhang würde das Datenschutzrecht oft als „Universalwaffe“ benutzt, obwohl es in vielen Belangen gar nicht um den Datenschutz ginge. Problematisch sei insbesondere der wenig definierte Begriff der „personenbezogenen Daten“. Aus dem Gesetz ergebe sich nicht, ob

darunter auch die massenhaft anfallenden Geodaten gehörten, kritisierte der Jurist.

In der anschließenden Podiumsdiskussion, die von Lena-Sophie Müller, der Geschäftsführerin der Initiative D21 e.V., moderiert wurde, stellte Ulrich Chiellino, Leiter Interessenvertretung Verkehr des ADAC e.V., die massenhafte Sammlung von Daten, die in selbst fahrenden Autos entstehen, und deren Sinn infrage. Eine Wahlfreiheit des Nutzers, selbst zu entscheiden, ob er automatisiert, eigenständig, online oder offline fahren will, sollte immer vorhanden sein. Dies forderte auch der Landesbeauftragte für den Datenschutz Baden-Württemberg Jörg Klingbeil und sprach sich dafür aus, dass datenfreies Fahren auch zukünftig möglich sein sollte. Seiner Ansicht nach dürften sich für Personen, die ihre Daten nicht weitergeben wollten,

keine Nachteile ergeben – etwa bei Versicherungstarifen. Ivo Körner, Geschäftsführer Vertrieb Branchenkunden IBM Deutschland GmbH, hielt dem entgegen, dass die Vernetzung bereits sehr weit vorangeschritten und eine Diskussion über Blackboxes im Auto nahezu hinfällig sei, wenn man berücksichtigen würde, dass fast jeder seine eigene Blackbox in Form eines Handys in der Hosentasche mit sich trage. Körner kritisierte zudem, dass Potenziale, die nicht nur einen wirtschaftlichen, sondern auch gesellschaftlichen Nutzen hätten, im Sinne des Datenschutzes oft nicht genutzt würden – wie etwa die Vernetzung zwischen Krankenwagen und Krankenhäusern bei Unfällen. Hier sollte eine intelligente Abwägung zwischen dem Schutz des Einzelnen und dem gesellschaftlichen Nutzen getroffen werden, so Körner. Zur finalen Frage, welche Weichenstellungen notwendig sind, um mit Big Data im Bereich Mobilität einen sinnvollen Umgang zu finden, gab es unterschiedliche Meinungen. Zukunftsforscher Mankowsky plädierte für eine amerikanisch-experimentelle, offenere Herangehensweise. Ein Mehrwert könne hier entstehen, wenn gleichzeitig die speziell deutschen Talente wie etwa technisches Ingenieurwissen, aber auch die Fähigkeit der tief-

Lena-Sophie Müller, Geschäftsführerin der Initiative D21 e.V., im Interview mit **Matthias Kammer**, Direktor, Deutsches Institut für Vertrauen und Sicherheit im Internet



Fotos: Stefan Zeitz Photography



Prof. Dr. Nikolaus Forgó, Institut für Rechtsinformatik, Leibniz Universität Hannover

gründigen Diskussion eingesetzt würden. Körner schloss sich dieser Meinung an und forderte eine Reduzierung des Datenschutzes in Bezug auf vernetzte Systeme, die einen Mehrwert für die Gesellschaft bieten. Der Landesdatenschutzbeauftragte Klingbeil verwies auf den Freiheitsgedanken,

der langfristig weiter eine Rolle spielen müsse. Wahlmöglichkeiten sollten erhalten bleiben, gleichzeitig aber Spielraum für Innovationen zugelassen werden. Chiellino sah im Vertrauen der Nutzer die wichtigste Weichenstellung für die Zukunft. Das Misstrauen bezüglich der Nutzung von Daten müsse

aktiv vermindert werden, indem die Nutzer einerseits mitgestalten können und andererseits gut informiert werden. Hier könnten auch technische Lösungen eine Rolle spielen.

Die Diskussion hat gezeigt, dass eine Vernetzung im Bereich der Mobilität durchaus wirtschaftlich und gesellschaftlich positive Effekte hervorbringen kann. Im Zentrum dieser Entwicklung sollte allerdings immer der Mensch selbst stehen, der intelligente Entscheidungen trifft und das Für und Wider abwägt – hier waren sich die Podiumsteilnehmer einig. Die Diskussion hat außerdem ergeben, dass bisherige rechtliche Regelungen wie etwa das Datenschutzrecht für viele Fragen der Vernetzung nicht zeitgemäß erscheinen, weshalb auch hier Innovationen erstrebenswert sind.



Alexander Mankowsky,
Zukunftsforscher Daimler AG



Ulrich Chiellino, Leiter Interessenvertretung Verkehr ADAC e.V.



2.3 Anwendungsfall Autoversicherung

In den vergangenen Jahren haben einige Autoversicherer begonnen, die Möglichkeit des Trackings bestimmter relevanter Daten in Autos zu nutzen, um neue Versicherungsmodelle anzubieten. Diese Varianten werden als telematikbasierte Versicherungen³² oder auch UBI (User Based Insurance) bezeichnet. Letztgenannter Begriff umschreibt allgemein alle Versicherungsmodelle, bei denen die Preisgestaltung vom Verhalten des Versicherten abhängt. Bezogen auf Kfz-Versicherungen werden auch die Ausdrücke „Pay as you drive“ oder „Pay how you drive“ verwendet. Traditionell werden die Prämien von Autoversicherungen anhand von allgemeinen historisch-statistisch ermittelten Faktoren berechnet, wie Alter, Geschlecht, Familienstand, Fahrzeugtyp, Wohnort, Gebrauchsinintensität, vorhergehende Inanspruchnahme der Versicherung und Scoring-Wert des Kunden. Die individuelle Fahrbilanz des Versicherten wird ebenfalls als ein Faktor berücksichtigt, allerdings beschränkt sich die Bewertung auf eine Aufstellung früherer Unfälle oder sonst nachträglich registrierter relevanter Vorkommnisse. Telematikbasierte Versicherungsmodelle hingegen versprechen, das tatsächliche Fahrverhalten des Versicherungsnehmers als vorrangigen Aspekt der Preisgestaltung zu beachten. Zu diesem Zweck wird bei heute bereits existierenden Modellen zumeist eine sogenannte Blackbox in das Auto des Kunden eingebaut, die den Fahrstil des Fahrzeughalters trackt. Eine Aufzeichnung der relevanten Daten ist aber auch über eine App im Smartphone des Kunden oder über Sensoren und Geräte möglich, die ohnehin ab Werk im Auto vorhanden sind.

Die Blackbox zeichnet das Verhalten des Fahrers unmittelbar und in Echtzeit auf. Die erfassten Daten wie die gefahrenen Kilometer, die Uhrzeit, der Ort, Geschwindigkeit, das Auslösen des Airbags, die Anzahl und Länge der Pausen bei langen Fahrten sowie Werte zu Beschleunigung und Bremsverhalten werden anschließend mittels einer eingebauten SIM-Karte versendet und vom Versicherer ausgewertet. Auf Basis der Datenanalyse werden schließlich die Prämien errechnet und entsprechend angepasst. Zumeist geschieht das, indem die Einzeldaten in verschiedene Scoring-Werte umgerechnet werden. So gibt es beispielsweise in einem konkreten Fall Punktabzüge für Geschwindigkeitsüberschreitungen, für Beschleunigungen mit mehr als 0,25g³³ oder Bremsungen mit mehr als 0,3g. Minuspunkte gibt es auch für Fahrten zwischen 23 und 6 Uhr. Die einzelnen Werte werden unterschiedlich gewichtet, und einmal im Monat wird der Gesamtscore errechnet, woraus sich dann der jeweilige Tarif ergibt.³⁴ Dabei finden die traditionellen Faktoren ebenfalls weiterhin Berücksichtigung. Darüber hinaus besteht für das verarbeitende Unternehmen die Möglichkeit, die Daten aller Kunden der jeweiligen Versicherung zu Big Data zusammenzuführen, um weiter gehende Erkenntnisse zu gewinnen.³⁵

Ein Vorreiter dieser Art von Autoversicherung ist das britische Unternehmen „Insure the Box“.³⁶ In Deutschland war es die Sparkassen-Direktversicherung, die Anfang Januar 2014 limitiert und befristet ein Telematikprodukt anbot.³⁷ Zuletzt kündigte der größte deutsche Autoversicherer HUK-Coburg im Mai 2015 an, künftig auch Versicherungstarife auf der Basis von Daten zu Geschwindigkeit sowie Brems- und Beschleunigungsverhalten aufzustellen.³⁸

32 Telematik steht für die Kombination aus Telekommunikation und Informatik; das Mittel der Informationsverknüpfung von mindestens zwei Informationssystemen mithilfe eines Telekommunikationssystems sowie einer speziellen Datenverarbeitung (Quelle: <http://de.wikipedia.org/wiki/Telematik>). Verkehrstelematik bezeichnet den Einsatz von Telematik im Verkehr. Sie beschäftigt sich mit der Erfassung, Verarbeitung und Darstellung von Daten zu Fahrzeugen mit ihrem zu transportierenden Inhalt, dem Versand und Empfang von Personen oder Gütern (Quelle: <http://de.wikipedia.org/wiki/Verkehrstelematik>).

33 „g“ ist hier die Erdbeschleunigung und eine Einheit, in der Beschleunigungs- und Bremsvorgänge quantifiziert werden können.

34 Kai Biermann, „Wer zu hart bremst, verliert seinen Versicherungsrabatt“, Zeit Online, 13. November 2013, online: <http://www.zeit.de/digital/datenschutz/2013-11/versicherung-telematik-ueberwachung-kfz/komplettansicht>.

35 Vgl. Bitkom, „Big Data und Geschäftsmodell-Innovationen in der Praxis: 40+ Beispiele“, Februar 2015, S. 91, online: <https://www.bitkom.org/Bitkom/Publikationen/Big-Data-und-Geschaeftsmodell-Innovationen-in-der-Praxis-40-Beispiele.html>.

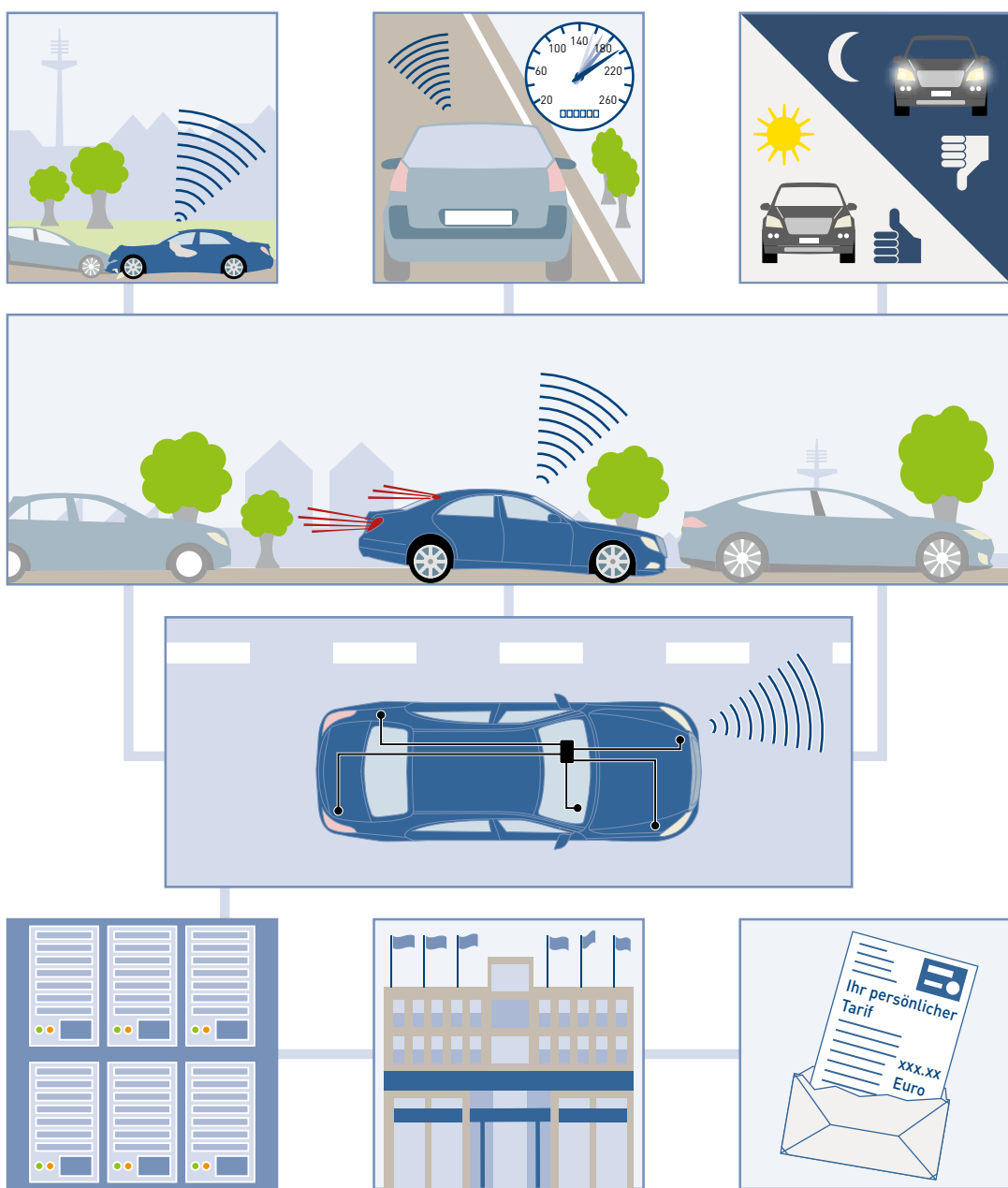
36 <https://www.insurethebox.com/about-insurethebox>.

37 <https://www.sparkassen-direkt.de/presse/telematik-sicherheits-service-startet.html>.

38 „HUK-Coburg plant Telematik-Tarife bei Autoversicherung“, Heise.de, 21. Mai 2015, online: <http://www.heise.de/autos/artikel/HUK-Coburg-plant-Telematik-Tarife-bei-Autoversicherung-2661601.html>.

TELEMATIK-BASIERTE AUTOVERSICHERUNG

Eine Blackbox im Auto erfasst Daten zum Fahrverhalten und übermittelt sie an die Versicherung. Die Auswertung dieser Daten bildet die Grundlage für die Höhe der Versicherungsprämie. Getrackt werden unter anderem starkes Bremsen oder Beschleunigen, Tageszeiten der Fahrten und zurückgelegte Entfernung.



„Die Differenzierung in Versicherungstarifen durch Auswertung individuellen Fahrverhaltens ist nicht gänzlich neu – sie gab es auch schon in der analogen Welt: Verursachen Sie einen Unfall, erhöht sich Ihr Tarif. Es stellt sich allerdings die Frage, wie detailliert und umfassend diese Differenzierung, die durch Big Data technisch möglich wird, werden soll.“

Jan Christian Sahl, Senior Manager Digitalisierung, Bundesverband der Deutschen Industrie e.V., Konsultation

Die Stakeholder

Im Rahmen dieser Interessenlage lassen sich folgende Stakeholder bezüglich telematikbasierter Autoversicherungen identifizieren:

- Die Versicherungen sind die offensichtlichsten Stakeholder dieser neuen Versicherungsart. So berichtet die Bitkom, dass es Insure the Box UK als erstem Autoversicherer gelungen ist, eine profitable trackingbasierte Kfz-Versicherung zu etablieren.³⁹ Zudem können über attraktive Preisgestaltungen, die durch telematikbasierte Modelle möglich sind, neue Kunden gewonnen werden. Den Angaben zufolge konnte Insure the Box UK im ersten Jahr nach Einführung der Versicherung 20.000 Neukunden gewinnen. Insofern ist die Interessenlage der Autoversicherer analog zu jener der Krankenversicherungen im Hinblick auf Gesundheitsdaten, die mittels Wearables generiert werden: Das Tracking der für die Versicherung relevanten Daten durch die Blackbox ermöglicht eine Preisdifferenzierung, die die statistischen Erhebungen traditioneller Versicherungsmodelle nicht zulassen.

- Autofahrer allgemein können Kosten sparen, wenn sie sich einen umsichtigeren und defensiveren Fahrstil angewöhnen. Darüber hinaus erhöht vorsichtigeres Fahren auch die eigene Sicherheit im Straßenverkehr.
 - Vor allem Versicherungskunden, die nach traditionellen Berechnungen stets im Vergleich deutlich höhere Prämien zu zahlen hatten, können von telematikbasierten Preisgestaltungen profitieren, da das Fahrverhalten unmittelbarer mit einbezogen werden kann. Das gilt insbesondere für Fahranfänger, die aufgrund statistischer Erhebungen über die Unfallzahlen junger Personen normalerweise bedeutend mehr für ihre Kfz-Versicherung zahlen müssen als erfahrenere Autofahrer. Dieser Nachteil kann zumindest teilweise ausgeglichen werden, wenn mittels einer Blackbox ein umsichtiger Fahrstil nachgewiesen werden kann.⁴⁰
 - Auch Gelegenheitsfahrer können Vorteile erlangen, da bei den meisten UBI-Modellen die Versicherungsprämie unmittelbar von den zurückgelegten Kilometern abhängt.⁴¹
- Ein besonderes Interesse könnten auch Verkehrsaufsichtsbehörden geltend machen, wenn die Akzeptanz des Versicherungsmodells dazu führt, dass sich die Sicherheit auf den Straßen erhöht.

„Datenfreies Fahren sollte bewusst weiter möglich sein. Menschen dürfen nicht dafür bestraft werden, wenn sie ihre Daten nicht hergeben wollen – etwa bei Versicherungstarifen.“

Jörg Klingbeil, Landesbeauftragter für den Datenschutz Baden-Württemberg, öffentliche Veranstaltung Stuttgart, 25.06.2015

³⁹ Bitkom, S. 91.

⁴⁰ Emma Lunn, „Car insurance ‘smartbox’ that could lower young drivers’ premiums“, The Guardian Online, 23. April 2011, online: <http://www.theguardian.com/money/2011/apr/23/car-insurance-smartbox-young-drivers>.

⁴¹ Vgl. <https://www.insurethebox.com/about-insurethebox>.

Die Chancen des Trackings bei Autoversicherungen

Die Chancen telematikbasierter Autoversicherungen ergeben sich unmittelbar aus den Interessen der Stakeholder. So können umsichtig fahrende Kunden erheblich Kosten sparen, und durch vorsichtiges Fahren reduziert sich zudem die Gefahr, dass sie in Unfälle verwickelt werden. Den Versicherern eröffnet sich die Möglichkeit, profitablere Produkte anzubieten. Der Autofahrer gewinnt insofern an Autonomie hinzu, als er unmittelbar Einfluss auf die Preise seiner Kfz-Versicherung nehmen kann. Außerdem erhöhen die neuen Preismodelle die Auswahl an verfügbaren Versicherungsprodukten. Weiterhin führt das Tracken des Fahrzeugs mittels der Blackbox dazu, dass es im Falle eines Diebstahls leichter auffindbar ist. Das Unternehmen Insure the Box UK beispielsweise reklamiert, seit Einführung des Modells mehr als 30.000 gestohlene Fahrzeuge ermittelt und den Diebstahl unzähliger weiterer Autos verhindert zu haben.⁴²

Neben den genannten Stakeholdern kann auch die Gesellschaft als Ganzes von der Einführung solcher Versicherungsmodelle profitieren. Dies betrifft zunächst die Verkehrssicherheit, denn von einer geringeren Unfallwahrscheinlichkeit profitieren nicht nur die individuellen Autofahrer, sondern auch andere Verkehrsteilnehmer. Berichte aus Italien scheinen den Zusammenhang zwischen der Verkehrssicherheit und telematikbasierten Versicherungsmodellen zu bestätigen. Laut eines Berichts der BBC ist dort die Unfallrate seit der umfassenden Einführung von Blackboxes um 16 Prozent gesunken.⁴³ Der gleiche Effekt könnte sich zudem daraus ergeben, dass Autos insgesamt weniger genutzt werden, wenn eine Verringerung der zurückgelegten Kilometer sich unmittelbar günstig auf die Versicherungskosten auswirkt. Geringere Nutzung könnte dann den Effekt haben, dass es zu weniger Unfällen und Staus kommt. Außerdem hätte dies zugleich eine Senkung der CO₂-Emissionen durch Autoverkehr und damit der Umweltbelastung zur Folge. Zu einer geringeren Belastung kommt es

sogar selbst dann, wenn zwar nicht weniger Auto gefahren, der Fahrstil insgesamt aber defensiver wird, die Autofahrer also weniger stark beschleunigen, im Durchschnitt langsamer fahren und damit weniger häufig scharf bremsen müssen.

„Individuelles Fehlverhalten kann auch mit Strafen sanktioniert werden. Dafür braucht man nicht zwingend Technik, die z.B. verhindert, dass man zu schnell fährt.“

Dr. Johannes Springer, Leiter Technologie und Geschäftsentwicklung, Konzerngeschäftsfeld Connected Car der T-Systems International GmbH, 2. Expertenworkshop, 15.06.2015

Sollte es dennoch zum Unfall kommen, ist die Beweisführung vereinfacht, wenn die relevanten Fahrdaten fortlaufend aufgezeichnet wurden. Das Risiko für ehrliche, vorsichtig fahrende Verkehrsteilnehmer, ihre Ansprüche aufgrund von Beweisschwierigkeiten nicht durchsetzen zu können, wird deutlich verringert. Die eingebauten Systeme können genutzt werden, um bei Unglücksfällen die Position des Fahrzeugs zu übermitteln und automatisch den Notruf zu wählen, so wie es ab diesem Jahr mit dem „eCall“-System für jedes in der Europäischen Union neu zugelassene Auto – aber eben nur für diese – verpflichtend ist. Schließlich können die Kunden davon profitieren, dass sie über Online-Plattformen der Versicherer ein konstantes Feedback über ihr Fahrverhalten bekommen, was als weiterer Anreiz wirken kann, ihr Fahrverhalten zu verbessern.

„Die Debatte über ein Thema wie Big Data sollte mehr mit Blick auf die Chancen als allein auf die Risiken geführt werden.“

Prof. Dr. Felix Sasaki, Senior Researcher Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI) und W3C Fellow, Konsultation

⁴² Vgl. Bitkom, S. 91.

⁴³ Eamonn Walsh & Vanessa Barford, „The proliferation of the little Blackbox“, BBC Online, 31. Oktober 2012, online: <http://www.bbc.com/news/magazine-20143969>.

Die Risiken des Trackings bei Autoversicherungen

Den genannten Chancen steht allerdings eine Reihe von problematischen Konstellationen gegenüber, die sich aus der Einführung telematikbasierter Autoversicherungen ergeben können. Die dringendsten Fragen stellen sich insoweit in Bezug auf den Schutz der Daten, die durch die in den Autos installierten Blackboxes erzeugt und an die Versicherungen übermittelt werden. Das gilt insbesondere für die Geolokationsdaten, die durch das GPS-Modul entstehen. Auf diese Weise können umfassende Bewegungsprofile der Autofahrer erstellt werden. Solche Ortsdaten sind schon deshalb als besonders sensibel einzustufen, weil sie sehr viele Rückschlüsse auf das Privatleben des Anwenders zulassen. Eine Studie konnte zeigen, dass nicht einmal die GPS-Daten selbst gespeichert werden müssen, um beispielsweise die Fahrtziele von Autofahrern, die mittels Blackbox getrackt werden, abzuleiten. Eine Big-Data-Analyse der sonstigen Daten zum Fahrverhalten wie Geschwindigkeit und die Reisezeit erwiesen sich als ausreichend.⁴⁴ Zum Problem wird dieser Umstand spätestens dann, wenn nicht mehr garantiert werden kann, dass die Daten ausschließlich für den Versicherungszweck verwendet werden.⁴⁵ Staatliche Behörden könnten versuchen, an die Daten zu gelangen, um einzelne Personen zu überwachen. Erinnert sei in diesem Zusammenhang an die Initiative des früheren Bundesinnenministers Hans-Peter Friedrich während der Koalitionsverhandlungen im Herbst 2013. Entgegen ursprünglich anderslautender Ankündigungen sollte die Gesetzeslage entsprechend geändert werden, damit Ermittler Zugang zu den Datensätzen aus

dem Autobahn-Mautsystem erhalten können.⁴⁶ Aber auch jenseits staatlicher Überwachung können die Geolokationsdaten von Interesse sein. In den Vereinigten Staaten beispielsweise finden GPS-Daten inzwischen sogar als Beweismittel in Scheidungsverfahren Eingang.⁴⁷

„Wahlfreiheit ist wichtig. Jeder sollte entscheiden können, ob er selbst fahren will oder automatisiert, online oder offline.“

Ulrich Chiellino, Leiter
Interessenvertretung Verkehr ADAC e.V.,
öffentliche Veranstaltung Stuttgart,
25.06.2015

Neben der Frage nach dem Datenschutz muss wiederum auch der Aspekt der Datensicherheit beachtet werden. Auch wenn Experten bislang nicht von einer erhöhten Gefährdung ausgehen: Klar dürfte sein, dass Autos desto mehr Angriffsflächen für potenzielle Hacker bieten, je mehr sie vernetzt sind.⁴⁸

Nicht unproblematisch ist auch die Struktur des Preismodells selbst. Es diskriminiert insofern, als manche Faktoren für das Scoring nichts mit dem eigentlichen Fahrverhalten zu tun haben: So werden u.U., in Abhängigkeit vom Versicherungstarif, sowohl nächtliche Fahrten als auch Fahrten in der Stadt negativ gewichtet. Wer in der Stadt wohnt und im Schichtdienst arbeitet, kann nicht oder jedenfalls nur selten von den günstigeren Tarifen profitieren.⁴⁹

Schließlich gilt für telematikbasierte Autoversicherungen das Gleiche, was schon im Hinblick auf diejenigen Krankenversicherungsmodelle diagnostiziert wurde, die Daten aus Fitness-Trackern für ihre Tarife heranziehen: Mit der Verantwortungsverlagerung weg von der Gesamtgesellschaft hin zum Indi-

⁴⁴ Rinku Dewri u.a., „Inferring Trip Destinations From Driving Habits Data“, Colorado Research Institute for Security and Privacy, Department of Computer Science, University of Denver, 2013, online: <http://www.cs.du.edu/~rdewri/data/MyPapers/Conferences/2013WPES-Extended.pdf>.

⁴⁵ In ihrer Antwort auf die Kleine Anfrage der FDP-Fraktion zur datenschutzrechtlichen Dimension von „Pay As You Drive“-Versicherungsmodellen vom März 2009 vertrat die Bundesregierung die Auffassung, die datenschutzrechtliche Problematik sei allein Sache der privatrechtlichen Vereinbarung zwischen Versicherung und Versicherungsnehmer. Bezüglich der Frage, wie gewährleistet werden solle, dass die Daten nicht auch für andere als den Versicherungszweck verwendet würden, beschränkte sich die Regierung auf die Aussage, ihr sei nicht bekannt, was die Versicherungswirtschaft diesbezüglich unternehmen wolle; siehe BT-Drucksache 16/12381 vom 23. März 2009, online: <http://dip21.bundestag.de/dip21/btd/16/123/1612381.pdf>.

⁴⁶ Vgl. Jörg Diehl u.a., „Überwachung: Innenminister Friedrich greift nach Maut-Daten“, Spiegel Online, 6. November 2013, online: <http://www.spiegel.de/politik/deutschland/ueberwachung-innenminister-friedrich-fordert-zugriff-auf-maut-daten-a-931952.html>.

⁴⁷ Siehe z.B. den Fall *Villanova v. Innovative Investigations, Inc.*, 420 N.J.Super. 353, 21 A.3d 650 (2011), online: http://privacylaw.proskauer.com/uploads/file/Villanova%20v_%20Innovative.pdf; vgl. dazu http://itlaw.wikia.com/wiki/Villanova_v._Innovative_Investigations.

⁴⁸ Vgl. Wilfried Eckl-Dorna, „Verräter“, Spiegel Wissen, 04/2014, S. 42.

⁴⁹ Kai Biermann, Wer zu hart bremst, s.o.

viduum geht eine Entsolidarisierung einher. Versicherungsmodelle, die am Verhalten des Kunden ansetzen, fördern so die Herausbildung neuer Normen und führen damit zu einem Anpassungsdruck. Schäden sind nicht mehr ein Ausdruck des allgemeinen Lebensrisikos, die deshalb von allen gemeinschaftlich zu tragen sind, sondern das Problem jedes einzelnen Betroffenen. Das bedeutet nicht nur einen Verlust individueller Freiheit. Für manche Personen, die beispielsweise aufgrund ihres Berufs nicht in den Anwendungsbereich der günstigen Tarife kommen können, kann das im Extremfall so weit führen, dass sie sich schlicht keine Kfz-Versicherung und damit eben auch kein Auto mehr leisten können.⁵⁰

„Big Data hat die Tendenz zur individuellen Zuschreibung von Verantwortung.“

Stephan Noller, Gründer und Geschäftsführer der ubirch GmbH,
2. Expertenworkshop, 15.06.2015

Herausforderungen für die Gesellschaft

Die genannten Chancen und Risiken des Einsatzes von Telematik bei Autoversicherungen verweisen auf die konkreteren Fragen, die auf gesellschaftlicher Ebene zu klären sind. So muss zunächst ein-

mal erörtert werden, inwieweit gefordert werden soll, dass die mittels der Blackbox getrackten Daten zweckgebunden nur durch den Versicherer selbst und nur im Rahmen des Vertrags verarbeitet werden dürfen. Daran anschließend ist zu fragen, wie sichergestellt werden kann, dass eine über den Vertragszweck hinausgehende Verwendung der Daten ausgeschlossen ist. Weiterhin besteht ein Risiko der Diskriminierung aufgrund von Eigenschaften, auf die die Betroffenen keinen Einfluss haben. Es muss insoweit geklärt werden, wie verhindert werden kann, dass bestimmte Bevölkerungsgruppen durch telematikbasierte Versicherungsmodelle systematisch benachteiligt werden.

„Wenn ich meine Daten für ein konkretes individualisiertes Angebot per Zustimmung freigebe, ist das in Ordnung. Es muss aber die Möglichkeit geben, dieses zu widerrufen. Außerdem dürfen die Daten nicht einfach von anderen Anbietern oder zu anderen Zwecken verwendet werden, sodass unkontrolliert ein Risiko für mich entsteht.“

Florian Schumacher, Gründer Quantified Self Deutschland und Digital Health Consultant iic-solutions, Konsultation

2.4 Anwendungsfall Verkehrsleitsysteme

Neben telematikbasierten Autoversicherungen gehören moderne, datengetriebene Verkehrslenkungsprojekte zu den exemplarischen Vorhaben im Bereich Smart Mobility, die sich im Verlauf des Projekts herauskristallisiert haben. Auch hier spielt Tracking zur Gewinnung von Daten eine entscheidende Rolle. Die Städte Stockholm und Toronto sollen im Folgenden als konkrete Anwendungsbeispiele dienen.

Das Beispiel Stockholm

Stockholm wird regelmäßig als eine der führenden Smart Cities in Europa genannt.⁵¹ Einen zentralen Pfeiler des städtischen Gesamtkonzepts stellt das Projekt zur Optimierung der Verkehrsströme in Stockholm dar, das 2010 von der Königlich Technischen Hochschule Stockholm in Kooperation mit IBM gestartet wurde. Die eingesetzte Anwendung ermöglicht die Zusammenführung, Auswertung und Verarbeitung von großen Datenmengen, die aus verschiedenen Quellen wie dem Tracking von GPS-Signalen, Kartenmaterial oder Daten von Wetterstationen gewonnen werden.⁵² Die Software kann damit sowohl strukturierte als auch unstrukturierte Datenströme in Echtzeit analysieren. Für die Verkehrslenkung in Stockholm wurden die Daten von GPS-Trackern verwendet, die in 1.500 Taxis und 400 Lastkraftwagen der Stadt angebracht worden waren. Auf diese Weise kamen innerhalb eines Jahres ca. 170 Millionen Datenpunkte zusammen.⁵³ Das Tracking von Fahrzeugen erfolgte nicht nur über eingebaute GPS-Module, sondern auch über Sensoren und Kameras an Straßen und Ampelanlagen. In Kombination mit den weiteren Quellen konnten die Verhaltensmuster der

motorisierten Verkehrsteilnehmer ermittelt werden, woraufhin sich Staus und andere Verkehrsprobleme zuverlässig prognostizieren ließen. Anschließend wurde es den Bürgern der Stadt ermöglicht, auf die Echtzeitinformationen über ihre Mobiltelefone zuzugreifen. Insgesamt konnte auf diese Weise das Verkehrsaufkommen um 20 Prozent gesenkt werden. Die durchschnittliche Fahrzeit von Pendlern wurde um 50 Prozent reduziert. Die CO₂-Emissionen verringerten sich zugleich um ungefähr 10 Prozent.⁵⁴ Nach dem Erfolg in Stockholm wurde das Konzept inzwischen von mehreren weiteren Städten übernommen, unter anderem London, Singapur und Brisbane.

Das Beispiel Toronto

Noch einen Schritt weiter ist man in Toronto gegangen. Die größte Stadt Kanadas kündigte im April 2015 an, ebenfalls ein über mehrere Jahre angelegtes Big-Data-Projekt zur Verkehrslenkung zu starten. Interessant an diesem Fallbeispiel ist insbesondere, dass es nicht ausschließlich um Stauvermeidung und den Autoverkehr gehen soll. Das Konzept ist umfassend angelegt und schließt auch den öffentlichen Personennahverkehr (ÖPNV) und den Radverkehr innerhalb der Stadtgrenzen mit ein. Damit soll es möglich werden, ein Gesamtbild der Fortbewegungsmuster in Toronto zu erhalten.⁵⁵ Die Daten aus den verschiedenen Kategorien der Fortbewegung sollen von unterschiedlichen Akteuren gesammelt und anschließend unter dem Dach des „Big Data Innovation Team“ zusammengeführt werden. Die Auswertung der Verkehrsdaten im Straßennetz soll ähnlich wie in Stockholm vor sich gehen, die Datenerhebung erfolgt softwaregestützt durch die McMaster University. Die Toronto Transit Commission ist für die Analyse der Bewegungs- und Verkehrsda-

51 Siehe z.B. Boyd Cohen, „The 10 Smartest Cities in Europe“, Fast Company Online, 13. Januar 2014, online: <http://www.fastcoexist.com/3024721/the-10-smartest-cities-in-europe>.

52 Alain Biem u.a., „IBM InfoSphere Streams for Scalable, Real-Time, Intelligent Transportation Services“, SIGMOD'10, 2010, S. 1093.

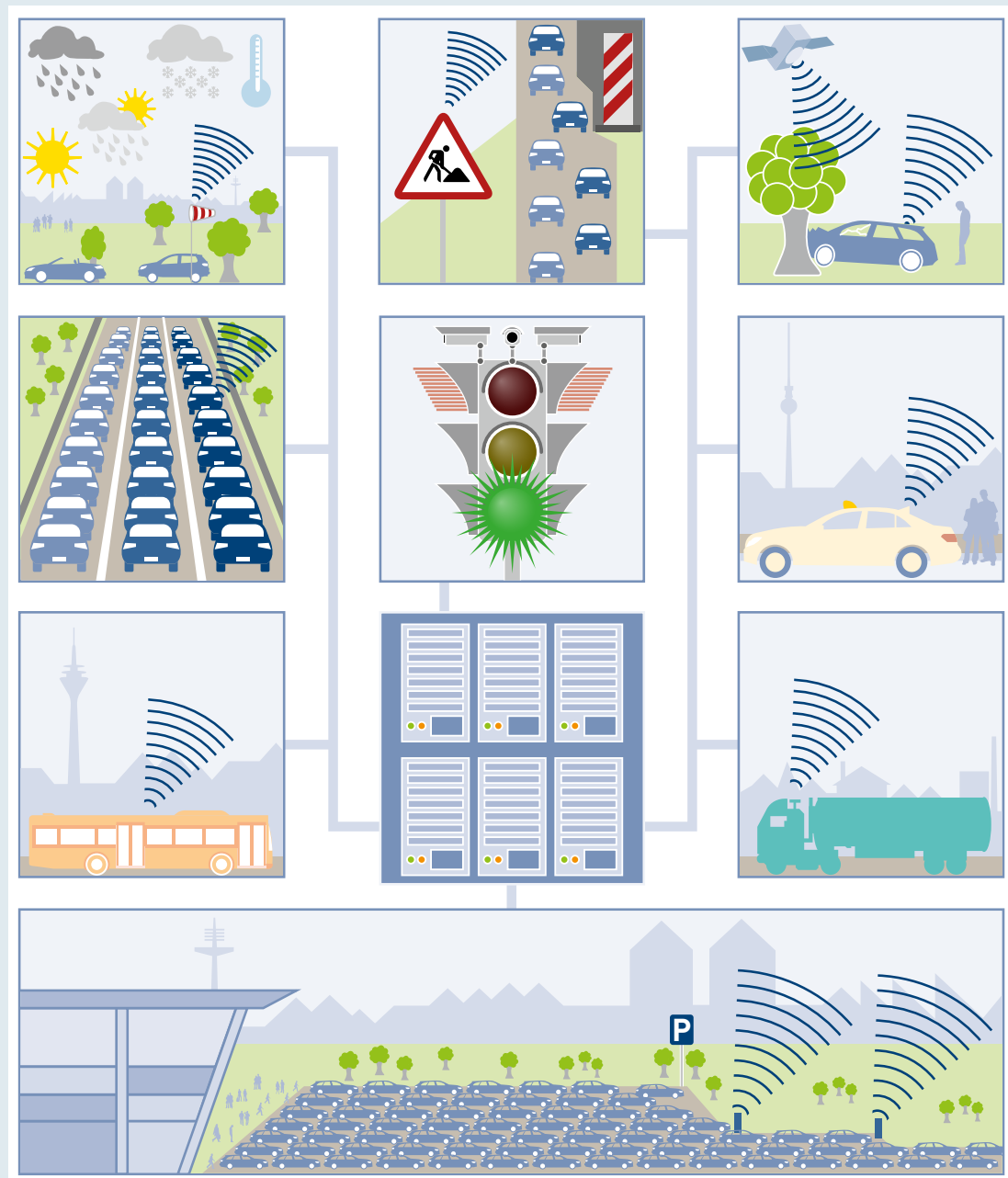
53 Alain Biem u.a., „Real-Time Traffic Information Management using Stream Computing“, Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 2010, S. 64, online: http://people.kth.se/~mahmoodr/Publications_files/A10JUN-CD2.pdf.

54 Hinzuzufügen ist allerdings, dass diese Ergebnisse auch deshalb erreicht werden konnten, weil ein Bestandteil des Gesamtprogramms zur Verkehrsleitplanung die Einführung einer Innenstadt-Maut war. Die somit gestiegenen Kosten des Individualverkehrs erhöhten den Anreiz, auf öffentliche Verkehrsmittel umzusteigen. Dieses „Congestion Charging System“ wird in einem Video von IBM erklärt: <https://youtu.be/rfMytzF4lv8>.

55 „City of Toronto to use traffic data to get Toronto moving“, Pressemitteilung, 7. April 2015, online: <http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=af71d7f79b2df6410VgnVCM10000071d60f89RCRD&nrkey=C17AD7EA49BB357685257E20004F9475>.

INTELLIGENTE VERKEHRSLLENKUNG

Sensoren und Tracker sammeln große Datenmengen – zu Verkehrsbewegungen, Unfällen, dem Wetter etc. – und übermitteln sie an Server, wo sie in Echtzeit ausgewertet werden. Aufgrund der Analyseergebnisse greifen Steuerungsmechanismen regulierend in den Verkehr ein und verhindern beispielsweise Staus.



ten des ÖPNV zuständig. Zu diesem Zweck könnte auf die Mobilfunkdaten der Nutzer des öffentlichen Nahverkehrs zurückgegriffen werden.⁵⁶ Für die Datenerhebung im Fahrradverkehr Torontos schließlich ist ein Tracking der Radfahrer mittels der Smartphone-Anwendung „Toronto Cycling App“ vorgesehen, die sich Freiwillige kostenfrei installieren können. Als Anreiz zur Teilnahme dient die App zugleich auch als Fitness-Tracker.⁵⁷ Da sich das Gesamtprojekt noch in der Startphase befindet, ist mit ersten Ergebnissen allerdings nicht vor dem nächsten Jahr zu rechnen.

Die Stakeholder

Bezüglich der Optimierung der Verkehrslenkung lassen sich die folgenden Stakeholder benennen:

- Die Verkehrsteilnehmer können am unmittelbarsten von den Maßnahmen Big-Data-gestützter Verkehrslenkung profitieren. Die Zeit, die sie mit Pendeln verbringen, kann im Idealfall signifikant verringert werden.
- Die Bürger haben insgesamt den Vorteil, in einer Stadt zu leben, in der Luft und Umwelt weniger durch Emissionen belastet sind.
- Durch eine umfassende Analyse der Pendlerdaten können die städtischen Verkehrsbetriebe in die Lage versetzt werden, den ÖPNV effizienter in Bezug auf Auslastung und Angebot zu betreiben, wodurch Kosten eingespart werden können. Zudem hat das Beispiel Stockholm gezeigt, dass die veröffentlichte Analyse der Daten des Individualverkehrs dazu geführt hat, dass mehr Bürger anschließend auf öffentliche Verkehrsmittel umgestiegen sind.
- Die Wirtschaft profitiert insofern, als einerseits Arbeitnehmer weniger Zeit im Verkehr verlieren. Andererseits können Waren effizienter durch die Stadt transportiert werden, wenn das Verkehrsaufkommen abnimmt oder jedenfalls optimiert wird.
- Die kommunalen Behörden können die Big-Data-Analysen für stadtplanerische oder sonstige städtische Projekte nutzen.

„Autonomes Fahren ist kein Selbstzweck, sondern macht Sinn für bestimmte Funktionen – wie etwa das Vermeiden von Unfällen und Klimaschäden.“

Minister für Verkehr und Infrastruktur Baden-Württemberg, **Winfried Hermann**, öffentliche Veranstaltung Stuttgart, 25.06.2015

Chancen und Risiken von Tracking bezüglich moderner Verkehrslenkung

Gerade im Angesicht der Prognosen zur voranschreitenden Urbanisierung der Weltgesellschaft in den kommenden Jahrzehnten scheint kaum ein Weg an der Optimierung des städtischen Verkehrsaufkommens vorbeizuführen. Dabei muss die Reduzierung insbesondere des Individualverkehrs mittelfristig als Ziel formuliert werden, um der Umweltbelastung durch Emissionen in Ballungsräumen Herr werden zu können. Moderne Verkehrslenkung kann insgesamt dazu beitragen, die vorhandenen Ressourcen effizienter zu nutzen. Richtungsweisend sind vor allem Ansätze wie der Torontos, sämtliche städtischen Verkehrsmittel in die Analyse mit einzubeziehen. Kurzfristig ergeben sich Vorteile für die Bürger, deren eigenes Zeitbudget durch optimierte Verkehrslenkung geschont wird. Das Tracking der einzelnen Fortbewegungsmittel eröffnet dabei die Möglichkeit, Erkenntnisse über das Verkehrsaufkommen in zuvor nicht gekannter Detailfülle zu gewinnen.

Das Tracking der Verkehrsteilnehmer zum Zwecke der Verkehrsdatenerhebung birgt dabei ähnliche Risiken wie Tracking mittels Blackbox im Auto. Zwar werden weniger verschiedene Einzeldaten für die Auswertung benötigt, und auch die Personalisierung der Daten ist von vornherein nicht notwendig, um den Zweck zu erreichen. Eine anonymisierte Erhebung ist hinreichend. Das heißt allerdings nicht, dass nicht unter Umständen doch Rückschlüsse auf die Identität bestimmter Per-

⁵⁶ „City embracing 'big data' to combat traffic congestion in Toronto“, CityNews, 7. April 2015, online: <http://www.citynews.ca/2015/04/07/city-embracing-big-data-to-combat-traffic-congestion-in-toronto/>; ein analoges Pilotprojekt der Verkehrs-Aktiengesellschaft Nürnberg zusammen mit einem Tochterunternehmen der Deutschen Telekom, das die anonymen Signalisierungsdaten aus dem Mobilfunknetz mit dem Ziel der Ableitung von Verkehrs- und Bewegungsströmen als Basis von Verkehrsanalysen auswerten sollte, ist kürzlich noch vor Beginn aufgrund von Datenschutzbedenken zurückgezogen worden; siehe: „VAG bedauert Irritationen und beendet Pilotprojekt mit Telekom-Tochter“, Pressemitteilung, 19. März 2015, online: <https://www.vag.de/presse/aktuelle-presseinfos/detail/article/vag-bedauert-irritationen-und-beendet-pilotprojekt-mit-telekom-tochter.html>.

⁵⁷ <http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=5c555cb1e7506410VgnVCM10000071d60f89RCRD>.

sonen gezogen werden können. Wie in allen Fällen von Big Data stellt sich hier wiederum das Problem der Aggregation. Es verbleibt auch bei dieser Art des Trackings zumindest die theoretische Möglichkeit der Überwachung individueller Verkehrsteilnehmer durch das Generieren personenbezogener Bewegungsdaten. Das gilt umso mehr, wenn Daten zusätzlich mittels im öffentlichen Raum installierter Kameras oder durch Auswertung der Aktivität von Mobilfunkzellen erhoben werden. Interessant ist insoweit wiederum das Beispiel Stockholm. Hier beschränkte sich das Tracking auf GPS-Module, die in Taxis und Lkws installiert waren. Die Autos von Privatpersonen wurden bewusst nicht mit einbezogen. Sobald ein Taxi Passagiere beförderte, wurden aus Datenschutzwägungen automatisch weniger Datenpunkte gesammelt, um die Gefahr der Entstehung von Bewegungsprofilen möglichst gering zu halten.⁵⁸ Trotzdem: Wenn sich vernetzte Autos wie erwartet erst einmal am Markt durchgesetzt haben und flächendeckend am Verkehr teilnehmen, dann ist davon auszugehen, dass die erforderlichen Daten hinsichtlich des Verkehrsaufkommens künftig direkt dort erhoben werden.⁵⁹

„Die Herausforderung bei Big Data ist, positive Nutzwirkungen zu entwickeln, ohne dass Schadenswirkungen entstehen.“

Dr. Bernhard Rohleder,
Hauptgeschäftsführer Bitkom e.V.,
1. Expertenworkshop, 26.02.2015

Je nach Ausgestaltung kann mit ressourcenoptimierten Verkehrskonzepten, die auf der Auswertung von Big Data beruhen, schließlich unter Umständen auch eine Einschränkung der individuellen Fortbewegungsfreiheit verbunden sein. Wie die Journalistin Katrin Elger treffend festgestellt hat, ist die Vision moderner Verkehrslenkung der „perfekte Fahrer, der alles daransetzt, ökooptimiert unterwegs zu sein“.⁶⁰

Das Freiheitsmoment der Unvernunft, das klassischerweise im Individualverkehr gerade in Deutschland immer mitschwingt, könnte unter dieser Prämisse über kurz oder lang abhandenkommen.

Herausforderungen für die Gesellschaft

Im Vergleich zu telematikbasierten Autoversicherungsmodellen ist die datengestützte Verkehrslenkung der deutlich weniger invasive Anwendungsfall von Big Data – trotz der soeben beschriebenen Gefahr einer Überwachung und Ausforschung der Individuen. Relevant ist allerdings die Frage der Anonymisierung. Wenn es für die Optimierung der Verkehrsströme nicht notwendig ist, personalisierte Daten zu erheben, dann muss im Vorhinein geklärt sein, wie sichergestellt werden kann, dass es nicht doch zu einer Identifizierung einzelner Personen aufgrund der verfügbaren Datensätze kommt. Die Möglichkeit, über die Aggregation verschiedener Datenquellen Rückschlüsse auf bestimmte Personen zu ziehen und damit spezifische Bewegungsprofile zu erstellen, muss als Problem adressiert werden. Das traditionelle Datenschutzrecht reicht hier gerade nicht aus, da es stets nur personenbezogene Daten behandelt, aber nicht für sich genommen anonyme Daten, die erst durch Aggregation zu personenbezogenen Daten werden.

„Big Data kann auch in vielen Fällen unter datenschutzrechtlichen Bedingungen unbedenklich genutzt werden – etwa durch technische Instrumente wie Anonymisierung.“

Jan Christian Sahl,
Senior Manager Digitalisierung,
Bundesverband der
Deutschen Industrie e.V., Konsultation

⁵⁸ Biem u.a., „Real-Time Traffic Information Management“, s.o., S. 66 f.

⁵⁹ Vgl. Errol S. van Engelen, „Big Data Analytics“, s.o.

⁶⁰ Katrin Elger, „Computer, übernehmen Sie“, Spiegel Wissen, 04/2014, S. 70.

3 Smart Health: Tracking und Wearables

Erneut sei ein Szenario vorangestellt, das sich in naher Zukunft so oder ähnlich abspielen könnte: *Sie fühlen sich schlecht, Sie haben das Gefühl, Ihr Kopf platzt gleich. Im Spiegel sehen Sie, dass Ihre Augen ganz rot sind, die Adern an Ihrer Stirn treten hervor. Und das nur wegen dieser paar Stufen vom Erdgeschoss in Ihr Büro im ersten Stock! Natürlich liegen die Medikamente für den Blutdruck mal wieder zu Hause. Es hilft nichts, Sie müssen kurz zum Arzt gegenüber, damit der Ihnen eine neue Packung verschreibt. Jetzt wieder nach Hause zu fahren, würde einfach zu lange dauern.*

Zum Glück dauert der Arztbesuch nur ein paar Minuten. Dank Ihres neuen Fitnessarmbands müssen Sie keine langen Formulare mehr ausfüllen, wenn Sie bei einem anderen Arzt als Ihrem Hausarzt sind. Alle Gesundheitsdaten, Blutdruck, Puls, Bewegungsdaten und Schlafphasen sind auf dem Armband gespeichert. Sie bekommen Ihr Rezept direkt mit, und beim Arzt wird Ihnen gesagt, dass Sie in Zukunft auch gleich in die Apotheke gehen können, wenn Sie Ihr Armband mit dem neuen Diagnoseserver verbinden. Dann ginge alles sogar noch schneller.

Natürlich sind Sie einverstanden, mit Ihrem Armband ans Netz zu gehen – zumal dadurch Ihre Versicherungsbeiträge gesenkt werden, weil die Versicherung Ihre Daten verwenden kann, um ihr Geschäftsmodell und ihren Leistungskatalog zu optimieren.

Seit ein paar Wochen sind Sie nun mit Ihrem Armband im Netz, und die Sache gefällt Ihnen sogar noch

besser als gedacht! Das Armband erinnert Sie an die Pillen für den Blutdruck, und die zugehörige App macht Ihnen Vorschläge, wann Sie etwas essen sollten – das vergisst man ja leicht an einem stressigen Tag im Büro – oder wann Sie sich mal einen Moment hinlegen sollten. Eine Analyse Ihrer Schlafdaten und weiterer Körperfunktionen hat ergeben, dass Sie mindestens eine halbe Stunde liegen sollten, um einen Erholungseffekt zu erreichen. Die zehnminütigen „Power naps“, die Sie früher gemacht haben, waren in Ihrem Fall also nichts anderes als Zeitverschwendung. Dank des Tracking-Armbands nutzen Sie Ihre Zeit heute sinnvoller. Und weil Sie die Ratschläge meistens befolgen, belohnt die Versicherung Sie mit weiteren Rabatten. Sie leben gesünder und sparen dabei noch Geld.

Weil Sie Ihre Daten auch für ein staatliches Forschungsprojekt freigegeben haben, tragen Sie außerdem dazu bei, dass die Therapiemöglichkeiten für Herzpatienten wie Sie besser weiterentwickelt werden können. Zweimal die Woche nehmen Sie jetzt an einer Lauf-Challenge für Herzpatienten teil, die das gleiche Armband tragen wie Sie. Da herrscht eine tolle Atmosphäre, und das Zugehörigkeitsgefühl wirkt zusätzlich motivierend.

Ein Bekannter aus der gleichen Laufgruppe erzählt Ihnen, dass er es einmal mit dem Laufen übertrieben hätte und Herzrasen bekam. Weil er ohne Handy joggen war, konnte er selbst keine Hilfe holen. Dadurch bekam er eine Angstattacke, was seinen Zustand noch verschlimmerte. Sein Tracking-Armband

hatte allerdings die beunruhigenden Körperwerte registriert und vorsorglich einen Arzt gerufen, der ihn dank der vom Armband übermittelten Geodaten leicht finden konnte. Ihr Bekannter ist der neuen Technik sehr dankbar, und auch Sie freuen sich über dieses Gefühl von Sicherheit.

Es geht Ihnen sehr gut, seit Sie die Ratschläge Ihres Armbands und der App befolgen. Sie haben viel mehr Energie im Alltag und fühlen sich als Herr über Ihren Körper. Nur Ihre Mutter ist etwas anstrengend. Sie versteht einfach nicht, dass Sie nun mittwochs keine Zeit mehr haben, mit ihr einkaufen zu gehen, weil Sie die nötigen Bewegungspunkte für Ihr Programm sonst nicht schaffen. Ständig hält sie Ihnen vor, Sie würden nur aus Gruppenzwang handeln! Sogar von Fremdbestimmung spricht sie.

An einem Tag im Büro bittet die Chefin alle Mitarbeiter, zusammenzukommen. Sie kündigt an, für die gesamte Firma Tracking-Armbänder anzuschaffen, um die Kosten für die betriebliche Krankenversicherung zu senken. Selbstverständlich seien die Armbänder für alle Mitarbeiter kostenlos. Bis zum Beginn des nächsten Monats sollen sich alle mit dem Server der betrieblichen Versicherung verbinden. Als die Chefin fragt, ob das für alle in Ordnung sei, gucken einige Kollegen verschämt zu Boden, aber niemand sagt irgendetwas.

Später werden Sie von einem Kollegen angesprochen. Er regt sich über die Neuerung auf. Konformitätsdruck sei das! Unfreiheit! Sie halten dagegen, dass das doch ein toller Service der Firma sei und schließlich doch nur dem Wohlbefinden aller diene. Er guckt nur verärgert und vertraut Ihnen ein Geheimnis an: Bei seinem vorigen Job im Ausland waren solche Armbänder schon gang und gäbe. Weil aber sein damaliger Chef Zugriff auf die Gesundheitsdaten der Mitarbeiter und die zugehörigen Analyse-Tools hatte, sei die Sache für ihn nicht besonders gut gelaufen. Die Software hatte erfasst, dass er immer Herzklopfen bekam, wenn die Frau seines damaligen Chefs im Büro war. Das war dann zwar nicht der offizielle Kündigungsgrund, aber die Andeutungen des Chefs waren eindeutig, als dieser ihm die Auswertung seiner Daten präsentiert hatte. Das sei doch gegen die Menschenwürde, so durchleuchtet zu werden! Sie antworten dem Kollegen, dass dies ein ganz klarer Missbrauch der neu-

en Technik sei. Danach haben Sie aber ein mulmiges Gefühl.

Sie ertappen sich dabei, sich überwacht zu fühlen. Der unzufriedene Kollege scheint ein Aufwiegler zu sein, und seine Probleme sind nicht die Ihren! Weil Sie wissen, dass die Armbänder über eine ziemlich genaue Ortungsfunktion verfügen, nehmen Sie sich vor, in Zukunft mit dem betreffenden Kollegen nicht mehr alleine zu reden, sondern lieber nur in der Gegenwart weiterer Kollegen (und deren Armbändern). Dann würden Ihre Daten auch keine falschen Schlüsse über Sie zulassen. Schließlich wissen Sie nicht, wie Ihre Daten interpretiert werden könnten.

Sie sind zu Hause. Die letzten Ereignisse gehen Ihnen nicht aus dem Kopf, Sie fühlen sich irgendwie unwohl, was auch Ihrem Armband nicht entgeht. Plötzlich klingelt eine Nachbarin und bittet Sie um Hilfe. Offenbar gab es einen Fehler mit ihrem Armband (denn auch ihr Arbeitgeber hat solche Armbänder angeschafft). Es behauptet, dass sie die letzten zehn Tage durchgehend gejoggt sei. Die Versicherung hat dies automatisch als Betrugsversuch gewertet, um Gesundheitsprämien zu erschleichen, und hat ihr sämtliche Bonusrabatte gestrichen. Weil die Nachbarin aber einen neuen „Fit-und-Vorteil“-Tarif abgeschlossen hatte, ist ihre Versicherung ohne die Tracking-Daten unbezahlbar. Eine andere Versicherung abzuschließen, ist ohne Tracking-Armband neuerdings jedoch nicht mehr möglich.

Sie erklären der Nachbarin, dass die Lösung dieses Problems wohl einige Zeit in Anspruch nehmen würde, während der ihr Versicherungsschutz auf ein Mindestmaß heruntergestuft würde. Solange geprüft würde, ob das Armband nur einen Fehler habe oder ob es manipuliert wurde, würden nur dringende medizinische Notwendigkeiten übernommen. Die Zahnspange für ihre Tochter gehört nicht dazu.

Die Situation Ihrer Nachbarin betrübt Sie zwar ein wenig, aber letztendlich ist bei dieser Person ohnehin Hopfen und Malz verloren. Als Sie ihr vor ein paar Wochen mit einem Update ihrer Gesundheits-App geholfen hatten, haben Sie gesehen, dass die Nachbarin sämtliche Warnmeldungen der Software ignoriert und entgegen den Vorschlägen nur wenig geschlafen hatte. Weil die Nachbarin auch ihr Einkaufsverhalten in die Gesundheitsanalyse mit einfließen lässt, konnten Sie sehen, dass sie eine starke Raucherin ist und sich

hauptsächlich von Fertigprodukten ernährt. In gewisser Weise ist sie also auch ohne den technischen Fehler selbst schuld, wenn ihre Versicherung jetzt zu teuer wird. Sie selbst leben schließlich viel gesünder. Warum sollten Sie über Ihre Beiträge solche Risikopersonen mitfinanzieren? Ingeheim sind Sie dennoch froh, dass Sie es sich leisten können, etwas mehr für Ihre Versicherung zu bezahlen, sodass Sie Ihre Einkaufsdaten nicht an den Server zu übermitteln brauchen. Denn ab und zu trinken Sie auch gerne ein Glas Wein mehr.

Als Sie mit Ihrer Mutter über die Probleme der Nachbarin sprechen, regt diese sich gleich wieder auf. Alles würde heutzutage gemessen! Sogar, was man isst!

Ihre Mutter kennt die Nachbarin gut und weiß, dass diese als Mutter von fünf Kindern einfach nicht besonders viel Schlaf bekommen könnte. Außerdem könne es sich nicht jeder leisten, Bio-Lebensmittel zu kaufen! Das würde aber nicht mitgemessen von unseren tollen Armbändern!

Sie kratzen sich am Kopf. Ihre Smart Watch zeigt gleichzeitig an, wie viele Kalorien Sie dabei verbrennen. Weil Sie aus unerfindlichen Gründen Herzklopfen haben, schlägt die App Ihnen einen Spaziergang an der Sonne vor.

3.1 Begriffsdefinition

Auch diese Geschichte hat so noch nicht stattgefunden. In gewisser Hinsicht ist sie aber noch näher an der Realität als jenes Szenario, das den Abschnitt über Smart Mobility einleitete. Denn Geräte zum Tracking von Gesundheitsdaten – sei es als Fitnessarmband oder als Smart Watch – haben sich in den vergangenen Jahren immer weiter verbreitet.

Tracking-Technologien spielen in der modernen Medizin eine zunehmend wichtige Rolle. Die fortlaufende Überwachung von Körperfunktionen findet vor dem Hintergrund einer weitreichenden Veränderung des Gesundheitssektors statt. Diese Umwälzung lässt sich unter dem Begriff „Smart Health“ zusammenfas-

sen. Damit wird in erster Linie umschrieben, dass der Datenerhebung und -verarbeitung in medizinischen Kontexten eine wachsende Bedeutung zukommt. Das Phänomen großer Datenmengen ist gerade auf diesem Gebiet alles andere als eine Neuigkeit – Daten waren schon immer Grundbausteine sowohl in der Forschung als auch bei der Diagnose und Behandlung von Krankheiten. Patientenberichte und Patientenakten, statistische Auswertungen über Verläufe von Epidemien, diagnostische Maßnahmen wie das Aufnehmen von Röntgenbildern, die Korrespondenz mit Patienten, Ärzten und Krankenhäusern bei den Krankenversicherungen – diese Daten spielten und spielen eine große Rolle in der medizinischen Forschung und Administration.

„Die Digitalisierung ist für ein modernes und effektives Gesundheitswesen zentral.“

Oliver Schenk, Abteilungsleiter G
„Grundsatzfragen der Gesundheitspolitik, Telematik“, Bundesministerium für Gesundheit, öffentliche Veranstaltung
Berlin, 30.09.2015

Durch die Digitalisierung sind neue Datenquellen hinzugekommen, die bislang nicht zugänglich oder überhaupt nicht vorhanden waren. So können Konversationen über Gesundheitsthemen in Sozialen Medien, Webforen und anderen Internet-Quellen für medizinpräventive Zwecke ausgewertet werden. Analysen des Einkaufsverhaltens von Kunden in Supermärkten können unter Umständen relevante Gesundheitsdaten generieren. Das bekannteste Beispiel hierfür ist die inzwischen wieder beendete Praxis der amerikanischen Supermarktkette Target, der es gelang, aus den Einkäufen registrierter Kundinnen festzustellen, ob diese schwanger waren, sodass sie gezielt Werbung verschicken konnte.⁶¹

Eine traditionelle Quelle für praktische medizinische Anwendungen sind wissenschaftliche Studien. Gerade hier hat das Datenvolumen längst eine Dimension erreicht, die das menschliche Auffassungsvermögen übersteigt. So wird beispielsweise geschätzt, dass es heute allein über Diabetes mehr als 400.000 Fachar-

⁶¹ Kashmir Hill, „How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did“, Forbes Online, 16. Februar 2012, online: <http://www.forbes.com/sites/kashmirhil/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/>.

tikel gibt.⁶² Es ist offensichtlich, dass selbst bestmöglich ausgebildete und sich stets auf dem neuesten Forschungsstand haltende Ärzte eine solche Informationsfülle weder aufnehmen noch verarbeiten können. Dies gilt für das theoretische Wissen in den Fachgebieten selbst, ist aber auch ganz konkret ein Problem für die praktische Anwendung. Um Menschen helfen zu können, müssen die Erkenntnisse aus der Forschung an die spezifischen Bedürfnisse des jeweiligen Patienten angepasst werden. Auch hier übersteigt die Datenmenge zunehmend menschliche Verarbeitungskapazitäten. Für das Jahr 2015 rechnet man damit, dass einzelne Patientenakten auf 20 Terabytes anwachsen werden. Für das nächste Jahrzehnt sind bereits Größen von 90 Zettabytes (90×10^{21}) vorausgesagt.⁶³ Ein solcher Reichtum an medizinisch relevanten Details eröffnet die Möglichkeit, Diagnose und Therapie auf die speziellen Bedürfnisse des einzelnen Patienten anzupassen. Es besteht die Hoffnung auf deutlich verbesserte, weil zielgerichtetere Heilungsansätze. Aber erst die Rechenkapazität und Software moderner Systeme wie beispielsweise diejenige des von IBM betriebenen Watson erlaubt es, die Bestände an Big Data sowohl auf Forschungs- als auch auf Patientenseite in hinreichender Geschwindigkeit und Präzision zu verarbeiten.⁶⁴

„Das größte Risiko, das derzeit besteht, ist, dass die Chancen der Digitalisierung und Vernetzung im Gesundheitswesen nicht gesehen werden.“

Dr. Franz Bartmann, Präsident der Ärztekammer Schleswig-Holstein und Vorsitzender des Ausschusses „Telematik“ der Bundesärztekammer, öffentliche Veranstaltung Berlin, 30.09.2015

Vor dem Hintergrund der jüngsten technischen Entwicklungen sind mit dem Einsatz von Big Data im Gesundheitsbereich große Chancen verbunden. So ist es

Ärzten beispielsweise inzwischen möglich, über Muster in großen Datensätzen zu erkennen, wie gewisse Therapien oder Medikamente in bestimmten Bevölkerungsgruppen wirken. Dadurch können Behandlungsmethoden besser an individuelle Bedürfnisse angepasst werden. Betreiber von Krankenhäusern können durch die Auswertung von Big Data ihre Ressourcen effizienter einsetzen und so Kosten sparen. Auch die Versicherungswirtschaft verspricht sich neue Wege, Gesundheitskosten zu senken, indem zum Beispiel durch Big-Data-Analysen Kostenprognosen für einzelne Versicherungsnehmer präziser erstellt werden können. Staatliche Gesundheitsbehörden wiederum könnten Big Data zum Zweck der Prävention oder zur Voraussage des Verlaufs von Epidemien nutzbar machen – auch wenn jüngere Analysen die Euphorie um Googles Big-Data-Tool zur Prognose von Grippewellen deutlich gedämpft haben, da sich erste erfolgreiche Anwendungen von „Google Flu“ nicht wiederholen ließen.⁶⁵ In einer neuen Studie konnte aber beispielsweise gezeigt werden, dass eine Auswertung von Twitter-Meldungen in Nigeria den Ausbruch der Ebola-Epidemie im Sommer 2014 drei bis sieben Tage vor den ersten offiziellen Verlautbarungen der örtlichen Behörden hätte nachweisen können.⁶⁶

„Der Fantasie sind eigentlich fast keine Grenzen gesetzt, wo man Big Data überall einsetzen kann. Wenn man Big Data ganz abstrakt versteht, bezogen auf den Einzelnen, kann es der Versuch sein, Verhaltensweisen vorauszuschauen bzw. Prognosen über Verhalten abzugeben.“

Dr. Silke Jandt, Vertreterin des Lehrstuhls für Öffentliches Recht, Informationstechnologierecht und Rechtsinformatik an der Universität Passau, Interview

⁶² Klaus Mainzer, „Big Data und die neue Weltordnung“, *viernull-magazin*, 02/2014, S. 29.

⁶³ Ebd.

⁶⁴ Vgl. <http://www.ibm.com/smarterplanet/us/en/ibmwatson/health/>; vgl. dazu Ben Kepes, „More Watson Goodness, IBM Announces Health-Focused Vertical“, *Forbes Online*, 16. April 2015, online: <http://www.forbes.com/sites/benkepess/2015/04/16/more-watson-goodness-ibm-announces-health-focused-vertical/>.

⁶⁵ Vgl. Charles Arthur, „Google Flu Trends is no longer good at predicting flu, scientists find“, *The Guardian Online*, 27. März 2014, online: <http://www.theguardian.com/technology/2014/mar/27/google-flu-trends-predicting-flu>.

⁶⁶ Michelle Odlum/Sunmoo Yoon, „What can we learn about Ebola outbreaks from tweets?“, *American Journal of Infection Control*, Vol. 43, No. 6, 2015, S. 563, online: <http://www.ajicjournal.org/article/S0196-6553%2815%2900137-6/fulltext>.

INTERVIEW MIT KARL MAX EINHÄUPL

Big Data – Chancen und Risiken im Gesundheitsbereich

? Die Chancen und Risiken von Big-Data-Anwendungen werden zunehmend diskutiert: Auf der einen Seite schnellere und genauere Diagnosen dank großer Datensätze, auf der anderen die Sorge um den gläsernen Patienten. Wie nehmen Sie in Ihrer Arbeit diese Debatte wahr? Können Sie im

Gesundheitsbereich eine Art Aufbruchsstimmung oder eher ein verstärktes Problembewusstsein ausmachen?

Karl Max Einhäupl: Die Digitalisierung macht wie überall im täglichen Leben auch nicht vor dem Gesundheitssektor halt. Unsere Bewegungen können vermessen werden, nahezu vollständige DNA-Sequen-

zierungen fordern uns heraus, unsere ethischen Regularien so zu definieren und zu konkretisieren, dass wir Missbrauch rechtzeitig erkennen und verhindern und gegebenenfalls strafrechtlich verfolgen können. Big Data ist auch im Zusammenhang mit der Globalisierung zu sehen. Hieraus können wir Chancen für alle Lebensbereiche ableiten, ohne dabei unser Handeln stets kritisch zu hinterfragen.

Prof. Dr. Karl Max Einhäupl

Karl Max Einhäupl ist seit 2008 Vorstandsvorsitzender der Charité – Universitätsmedizin Berlin. Nach dem Medizinstudium und der Promotion an der Ludwig-Maximilians-Universität absolvierte er von 1977 bis 1981 seine klinisch neurologische und psychiatrische Ausbildung am Klinikum Großhadern in München. 1986 habilitierte sich Pro-

fessor Einhäupl für das Fachgebiet Neurologie, 1988 wurde er zum Professor an die Neurologische Klinik der LMU München mit dem Schwerpunkt „Neurologische Intensivmedizin“ berufen. Vier Jahre später folgte der Ruf an die Humboldt-Universität Berlin. Von 1993 bis 2008 war Professor Einhäupl Direktor der Klinik für Neurologie der Charité, von

Foto: Charité



Der Datenschutz ist ein Grundrechtsschutz. Privatheit und die Mitgestaltung hinsichtlich der Preisgabe und auch der Verwendung der Daten des Einzelnen dürfen deshalb keine untergeordnete Rolle spielen.

Karl Max Einhäupl

? Während in vielen Bereichen neue Datenquellen erschlossen werden (zum Beispiel Geolokalisierung von Fahrzeugen), sind große Datensätze von jeher Teil der medizinischen Forschung. Was ist neu

angesichts der erweiterten Möglichkeiten durch Tracking und Big Data? Was hat sich im Gesundheitsbereich im Vergleich zu den letzten Jahrzehnten geändert?

KME: Neu ist, dass wir nun die Möglichkeiten haben, Daten besser zu erfassen, zum Beispiel durch den Einsatz von Sensoren zur Messung von Puls, Blutdruck, Blutzucker – und diese in Echtzeit mit der elektronischen Akte eines Patienten im Klinikum verbinden können. Die Verarbeitung der Daten muss nun verstärkt angegangen werden, um Ärzten echte Werkzeuge an die Hand geben zu können, die sie in ihren Entscheidungen unterstützen können.

Durch die Erfassung digitaler Gesundheitsdaten ist es möglich geworden, medizinisch-wissenschaftliche Fragestellungen schneller und effizienter zu erfor-

schen und herkömmliche Therapieformen zu hinterfragen und zu optimieren. Dafür brauchen wir aber leistungsfähige und nicht nur komplexe Datenbanken. Wir sind heute weitaus besser und schneller in der Lage, Gesamtzusammenhänge zu erkennen und Therapieformen umzustellen.

? **Wie schätzen Sie die Risiken ein, die sich daraus ergeben, dass Profile über Einzelpersonen mithilfe von Sammlungs-, Tracking- und Analysetechniken erstellt werden können?**

KME: Bei all diesen Themen darf die Selbstbestimmung des Einzelnen nicht außer Acht gelassen werden. Der Datenschutz ist ein Grundrechtsschutz. Privatheit und die Mitgestaltung hinsichtlich der Preisgabe und auch der Verwendung der Daten des Einzelnen dür-

2007 bis 2008 Leiter des Centrums für Neurologie, Neurochirurgie und Psychiatrie. Der Neurologe war von 2001 bis 2005 Vorsitzender des Wissenschaftsrates und ist Mitglied der Leopoldina. 2004 wurde ihm das Bundesverdienstkreuz verliehen.



BIG DATA – CHANCEN UND RISIKEN IM GESUNDHEITSBEREICH



fen deshalb keine untergeordnete Rolle spielen. Es geht um den Wert des Menschen und nicht um den Wert seiner Daten. Es darf keine ideologische Ungleichwertigkeit geben.

? **Wie lässt sich eine potenzielle Diskriminierung von Einzelpersonen oder Personengruppen vermeiden, die auf Big Data zurückzuführen ist?**

KME: Die Big-Data-Technik wird von Google, Facebook, Amazon, aber auch von den Finanzbehörden und der Polizei genutzt. Die technischen Möglichkeiten müssen dabei so ausgestaltet werden, dass die datenschutzrechtlichen Grundsätze der Zweckbindung und der Datensparsamkeit in einem ausgewogenen Verhältnis zueinander stehen. Handlungsbedarf besteht in der Anpassung der entsprechenden gesetzlichen Grundlagen. In der transnationalen Forschung erfolgt die Verarbeitung der patientenbezogenen Daten in der Regel pseudonymisiert beziehungsweise anonymisiert, und durch ein sogenanntes Consent Management auch nur durch die Zustimmung des Patienten.

? **Wie sehen Sie den Unterschied zwischen der Datensammlung eines staatlichen Krankenhauses, einer Krankenkasse oder der Datensammlung von Facebook? Welche Chancen und Risiken machen Sie diesbezüglich aus? Sollte es verschiedene Regeln für verschiedene datenverarbeitende Stellen geben?**

KME: Patientendaten unterliegen durch die strengen Regelungen der Landeskrankengesetze (LKHG) der Länder klaren Regularien. Auch durch die entsprechenden Berufsordnungen wird ein gesetzeskonformes Verhalten des Arztes und seiner Gehilfen vorausgesetzt mit dem Risiko der persönlichen strafrechtlichen Haftung. Wenn Krankenhäuser oder Krankenkassen Patientendaten erheben und verarbeiten wollen, bedarf es immer einer Rechtsgrundlage oder einer Einwilligung. Wir haben derzeit noch das Problem, dass wir keine einheitliche europäische Regelung haben. Hier wird hoffentlich die EU-Datenschutz-Grundverordnung, die sich in der Abstimmung der Mitgliedstaaten befindet, Klarheit bringen.

Wenn ich mich bei Facebook oder anderen Sozialen Netzwerken mit Nutzungsbedingungen einverstanden erkläre, verstehe ich in der Regel doch gar nicht, was sich dahinter verbirgt. Wie in der Medizin sollte es eigentlich selbstverständlich sein, dass Einwilligungen nur dann abgegeben werden können, wenn aufgeklärt wurde, und zwar in verständlicher Form.

? **Welche Daten sollten der Wirtschaft vorenthalten werden? Gibt es Daten, die Sie für zu sensibel halten, um sie für die Nutzung durch Big-Data-Anwendungen zu öffnen? Zählen Gesundheitsdaten dazu? Oder sollte gerade hier der Zugang geöffnet werden, um privatwirtschaftliche Forschung zu fördern?**

KME: Gesundheitsdaten sind besonders sensible Daten und unterliegen wie alle personenbezogenen Daten dem Grundsatz der Zweckbindung. Medizinische Forschung ist eine universitäre Aufgabe und keine Aufgabe für die Privatwirtschaft. Wenn Versicherungen überlegen, Boni einzuführen, indem Versicherungs-

Big Data ist in Bereichen von Vorteil, in denen es um die Therapie und Verbesserung der Heilungschancen von schweren Krankheiten geht. **Karl Max Einhäupl**

nehmer ihre personenbezogenen Gesundheitsdaten sammeln und an die Versicherung übermitteln, muss das kritisch betrachtet werden. Was auf den ersten Blick nach einem persönlichen Vorteil aussieht, kann am Ende ein Verlust sein, etwa wenn doch eine schwere Erkrankung eintritt. Wornach soll der Wert von Gesundheitsdaten eingeschätzt werden? Diese werden im Zweifel einen höheren Wert haben.

? **Big Data birgt nicht nur Risiken, sondern bietet sowohl der Gesellschaft wie auch dem Individuum viele Vorteile – gerade auf dem Gebiet der medizinischen Forschung. Können Sie hierzu Beispiele nennen, die Sie besonders eindrucksvoll finden?**

KME: Big Data ist in Bereichen von Vorteil, in denen es um die Therapie und Verbesserung der Heilungschancen von schwe-

ren Krankheiten geht, zum Beispiel in der Krebsforschung, in der Zusammenführung von sogenannten Biobanken, in denen biologisches Material von Patienten gesammelt wird. Weitere Beispiele solcher Bereiche sind das Berliner Institut für Gesundheitsforschung (BIH – Berlin Institute of Health), das verschiedene medizinische Disziplinen miteinander verknüpft, oder auch die Nationale Kohorte, die in den nächsten zehn Jahren 200.000 Menschen zwischen 20 und 69 medizinisch untersucht und die gewonnenen Daten für die detaillierte Erforschung von Volkskrankheiten wie Diabetes, Herz-Kreislauf-Krankheiten oder Krebs verwendet.

? **Wie kann die Nutzung von Big-Data-Analysen die Diagnosestellung und generell die Arbeitsweise von Medizinern verändern? Welche Veränderungen beobachten Sie bereits?**

KME: Durch die Nutzung hochgradiger technischer Geräte in der Medizin, die daraus resultierenden medizinischen Befunde, die kontinuierliche Erfassung von Krankheitsverläufen in Registern bis hin zur Abrechnung mit den Krankenkassen wird eine Vielfalt von Daten produziert, die es uns ermöglicht, Patienten auf einem immer steigenden Niveau zu behandeln und die Behandlungsdauer zu verkürzen. Wir gewinnen dadurch die Möglichkeit, für das Krankenhausunternehmen richtungsweisende unternehmerische Entscheidungen zu treffen, auch im Hinblick auf Spezialisierungen in der Krankenhausbehandlung.

? **Was wünschen Sie sich ganz persönlich von den Beteiligten in der Debatte um die Ausgestaltung und Nutzung von Big Data? Brauchen wir mehr Geschwindigkeit, mehr Ruhe oder auch ganz andere Blickwinkel?**



BIG DATA – CHANCEN UND RISIKEN IM GESUNDHEITSBEREICH

» **Medizinische Apps können für chronische Patienten durchaus hilfreich sein, wenn sie gezielt zur Kontrolle von Symptomen und Therapien eingesetzt werden.** *Karl Max Einhäupl*



KME: Wir brauchen zum einen mehr Geschwindigkeit, um in der Gesundheitsforschung nicht von Unternehmen wie Google oder Apple überholt zu werden. Dafür müssen die gesetzlichen, aber auch die tatsächlichen Voraussetzungen geschaffen werden. Die Chancen, die sich uns eröffnen, sind geeignet, Lebensverhältnisse zu verändern und anzupassen. Wir können derzeit Daten, die wir zweckgebunden erhoben haben, nicht ohne Weiteres in einem anderen Kontext nutzen. Wir benötigen dafür einen ethischen und gesellschaftlichen Konsens.

? **Wie bewerten Sie in Ihrer Funktion als Vorstandsvorsitzender der Charité-Universitätsmedizin Berlin die zunehmende Verbreitung von Fitnessarmbändern und Gesundheits-Apps?**

KME: Durch Fitnessarmbänder und Apps hat der Einzelne erstmalig die Möglichkeit, ohne Arztbesuch Infor-

mationen über seinen Körper aufzuzeichnen, zu analysieren und den eigenen Lebensstil zu hinterfragen, zum Beispiel „Wie viele Schritte bin ich heute gelaufen?“. Insofern können diese Hilfsmittel schon motivierend sein, mehr Verantwortung für sich selbst zu übernehmen. Medizinische Apps können für chronische Patienten durchaus hilfreich sein, wenn sie gezielt zur Kontrolle von Symptomen und Therapien eingesetzt werden. Vor allem der Bereich der Telemedizin wird zukünftig stark davon profitieren. Wenn die Daten sicher übertragen werden und deren Qualität gewährleistet ist, bietet dies viele Chancen für klinische Prozesse der Zukunft.

? **Sehen Sie es grundsätzlich als Vorteil oder Nachteil, dass Gesundheitsdaten wie etwa die tägliche körperliche Aktivität oder die Herzaktivität über die Nutzer bekannt sind?**

KME: Welche Daten der Einzelne bekannt gibt oder nicht, muss ihm

selbst überlassen bleiben. Mit der zunehmenden Digitalisierung wird sich der Einzelne mit dieser Fragestellung auseinandersetzen müssen. Das hängt auch damit zusammen, wie aufgeklärt der Einzelne darüber ist, welche Folgen die freizügige Freigabe von persönlichen Daten hat. Hier lässt sich durchaus gesellschaftspolitischer Handlungsbedarf erkennen.

? **Finden Sie es in Ordnung, wenn Versicherungen diese Daten auswerten können und die Ergebnisse mit ihren Vertragskonditionen verknüpfen?**

KME: Die Analyse von anonymen Gesundheitsdaten könnte man wohl mit Marktforschung gleichsetzen. Bei detaillierten Kenntnissen wird aber die Daseinsvorsorge an Bedingungen geknüpft, was kritisch ist. Die Teilnahme an solchen Programmen ist zwar freiwillig, aber verbunden mit sozialem und gegebenenfalls auch beruflichem Druck.



3.2 Smart Health als Anwendungsfeld für Tracking

Tracking im Gesundheitswesen

Da IT-Systeme heutzutage in der Lage sind, unvorstellbar große Mengen an Datensätzen zuverlässig und schnell zu verarbeiten und auszuwerten, hängt es in erster Linie von der Anzahl erzeugter oder verfügbarer, medizinisch relevanter Einzeldaten ab, ob es möglich ist, die gesundheitliche Versorgung immer weiter individualisieren und damit verbessern zu können, so die Annahme der Befürworter des Einsatzes von Big Data. Hier kommt Tracking als jüngste Datenquelle im Gesundheitssektor ins Spiel: Wenn es gelänge, von einer möglichst großen Anzahl von Personen gesundheitsrelevante Daten fortlaufend und über einen großen Zeitraum zu sammeln, dann würde die zur Verfügung stehende Datenbasis hinreichend anwachsen, um die beschriebenen Ziele zu erreichen. Und eben dies wird künftig flächendeckend möglich sein – dank der Einführung von Geräten, mit denen ein Nutzer selbstständig Einzeldaten aufzeichnen und anschließend auf die Server der Anbieter übertragen kann. Zu diesen Einzeldaten zählen etwa die Ernährung (z.B. eingenommene Kalorienmenge, Art der Nahrung), der körperliche Zustand (Blutdruck, Puls, Körpergewicht, Body-Mass-Index, Blutzucker, Körpertemperatur, Hirnaktivität, Atemfrequenz), der emotionale Zustand (Messung von Erregungszuständen) und die körperliche Tätigkeit (Schrittzähler, Überwachung der Schlafdauer und der Schlafphasen). Die auf den Servern der Anbieter zusammenlaufenden Daten der einzelnen Anwender können anschließend zu Big Data aggregiert und weiterverwendet werden.

Die Webseite der Quantified-Self-Bewegung, die als Vorreiter in Sachen Activity Tracking gilt und schon 2007 von den beiden amerikanischen Journalisten Gary Wolf und Kevin Kelly gegründet wurde, listet inzwischen mehr als 500 Tools auf, um die eigenen Aktivitäten zu tracken.⁶⁷ Neben traditionellen, analogen Metho-

den der Aufzeichnung von Körper- und Aktivitätsdaten in Tagebüchern oder Tabellen dominieren heute vernetzte Geräte, die die Daten automatisiert aufzeichnen. Dazu gehören zum Beispiel ans Internet angeschlossene Waagen, die Körperfett und Body-Mass-Index berechnen, oder in die Matratze integrierte Schlafsensoren. Insbesondere die Smartphones der jüngsten Generation haben zu einem Anstieg der Tracking-Aktivitäten und damit der Menge an verfügbaren Daten geführt. Schätzungen zufolge gibt es inzwischen mehr als 100.000 verschiedene Gesundheits-Apps für mobile Geräte, mit denen Nutzer ihre eigenen Werte tracken können.⁶⁸ Die auf Apples iPhone vorinstallierte App „Health“ beispielsweise zeichnet die Gesundheitsdaten des Besitzers fortlaufend auf, ohne dass dieser das Programm selbst aktivieren muss. Dabei werden Daten wie die Schrittzahl, die zurückgelegte Wegstrecke oder die Anzahl der bewältigten Treppenstufen anhand des im Smartphone eingebauten Beschleunigungssensors, des Gyroskops sowie des GPS-Chips ermittelt. Obwohl die App nicht deinstalliert werden kann, sichert das Unternehmen zu, dass die Daten beim Nutzer verbleiben und nicht vom Unternehmen selbst verwendet werden.⁶⁹ Als Reaktion auf Proteste von Nutzern hat Apple mit dem neuesten Update des Betriebssystems (iOS 8.2) allerdings die Möglichkeit implementiert, das Activity Tracking zu deaktivieren.⁷⁰

„Sehr viele Big-Data-Anwendungen oder Algorithmen haben das Ziel, Verhaltensänderung zu motivieren. Fraglich ist, wann so eine Verhaltensänderung erwünscht oder erzwungen ist und ab wann sie als negativ oder positiv empfunden wird.“

Prof. Dr. Sebastian Spaeth, Professor für Betriebswirtschaftslehre und Digitale Märkte an der Universität Hamburg, 1. Expertenworkshop, 26.02.2015

⁶⁷ <http://quantifiedself.com/>.

⁶⁸ Deborah Lupton, „Digital Health Technologies and Digital Data: New Ways of Monitoring, Measuring and Commodifying Human Embodiment, Health and Illness“, Olleros u.a. (Hg.), *Research Handbook on Digital Transformations*, Northampton 2016, S. 6, online: <http://ssrn.com/abstract=2552998>.

⁶⁹ Matthias Schüssler, „Auf Schritt und Tritt überwacht von Tim Cook?“, *Tagesanzeiger Online*, 10. November 2014, online: <http://www.tagesanzeiger.ch/digital/kummerbox/Auf-Schritt-und-Tritt-ueberwacht-von-Tim-Cook/story/28089199>.

⁷⁰ Michael Keller, „Mit iOS 8.2 könnt Ihr den Schrittzähler deaktivieren“, *Curved*, 15. Dezember 2014, online: <https://curved.de/news/mit-ios-8-2-koennt-ihr-den-schrittzaeher-deaktivieren-185167>.

Öffentliche Veranstaltung in Berlin

TREUER ASSISTENT ODER TROJANER AM KÖRPER? – WIE GESUNDHEITSTRACKING UNSEREN ALLTAG VERÄNDERT

Im Rahmen der sechsten öffentlichen Veranstaltung im Projekt „Braucht Deutschland einen Digitalen Kodex?“, die am 30. September 2015 im Berliner Meistersaal stattfand, diskutierten Experten aus Wirtschaft, Politik, Medizin und Wissenschaft über den Einsatz von Big Data im Gesundheitswesen. Ein besonderer Schwerpunkt lag dabei auf dem Thema Tracking mit Wearables und den damit verbundenen gesellschaftlichen Chancen und Risiken.

In seinem Grußwort wies Oliver Schenk, Abteilungsleiter G „Grundsatzfragen der Gesund-

heitspolitik, Telematik“ des Bundesministeriums für Gesundheit, zunächst auf das große Potenzial einer stärkeren Vernetzung der Gesundheitsdaten hin. Der dadurch erhöhte Informationsaustausch würde die Behandlungsmethoden und damit auch die Heilungsaussichten der Patienten verbessern. Eine umfassende Digitalisierung im Bereich Gesundheit sei daher auch ein wichtiges politisches Ziel auf dem Weg zu einem effizienteren und moderneren Gesundheitssystem, so Schenk. Die Probleme, die durch die Vernetzung zum Beispiel im

Bereich des Datenschutzes entstünden, müssten im Dialog mit allen Interessengruppen gelöst werden. Auch das Bundesministerium für Gesundheit arbeite daran aktiv mit. Hier nannte Schenk das E-Health-Gesetz, das die nötigen Voraussetzungen für eine sichere digitale Kommunikation im Gesundheitswesen schaffe.

Unter dem Titel „Big Data in der medizinischen Praxis – Die Zukunft hat bereits begonnen“ stellte Dr. Franz Bartmann in seiner Keynote heraus, dass die Nutzung von Gesundheitsdaten in vielen Bereichen der Medizin große Vorteile hervorbringen könnte. So wäre es möglich, über bestimmte Datenzusammenhänge neue Erkenntnisse über die Entstehung von Krankheiten zu ermitteln und zudem die individuelle Behandlung von Patienten zu verbessern. Der Präsident der Ärztekammer Schleswig-Holstein und Vorsitzende des Ausschusses „Telematik“ der Bundesärztekammer kritisierte die allgemeine Skepsis und Unsicherheit vieler Ärzte gegenüber der Vernetzung von Gesundheitsdaten. Das größte Risiko bestehe darin, dass diese Abwehr



Fotos: Tom Maelsa

Lena-Sophie Müller, Geschäftsführerin der Initiative D21 e.V., im Interview mit **Peter Schaar**, Vorstand der Europäischen Akademie für Informationsfreiheit und Datenschutz (EAID)

dazu führe, dass die Chancen, die daraus entstehen könnten, nicht wahrgenommen würden. Dr. Bartmann wies allerdings auch darauf hin, dass die Nutzung von Big Data in der Medizin die Gesellschaft vor ethische Herausforderungen stellen würde. So könnten durch bestimmte Muster in Big-Data-Analysen Wahrscheinlichkeitsaus-

sagen getroffen werden, die für ausgewählte Gruppen Krankheiten vorhersagen. Diese würden eventuell zwar gar nicht auftreten, die Vorhersagen hätten aber dennoch negativen Einfluss auf diesen Personenkreis – etwa durch Ängste oder Verhaltensänderung. Hier sei ein gesellschaftlicher Dialog über die Grenzen der Big-Data-Analy-

sen, Korrelationsaussagen und Prognosen notwendig, so Dr. Bartmann.

Ein Ring, der beim Eingang von E-Mails am Finger vibriert, ein Babybody mit Sensor, der vor dem plötzlichen Kindstod warnt, ein Wearable als Verhütungsmethode – in seiner Keynote stellte Stephan Noller, Gründer und Geschäftsführer



Stephan Noller, Gründer und Geschäftsführer der ubirch GmbH



Podiumsdiskussion: (v.l.n.r.) **Dr. Franz Bartmann**, Präsident der Ärztekammer Schleswig-Holstein und Vorsitzender des Ausschusses „Telematik“ der Bundesärztekammer; **Prof. Dr. Dagmar Borchers**, Professorin für Angewandte Philosophie an der Universität Bremen; **Peter Schaar**, Vorstand der Europäischen Akademie für Informationsfreiheit und Datenschutz (EAID); **Lena-Sophie Müller**, Geschäftsführerin der Initiative D21 e.V.; **Burkhard Dümler**, Director Development Digital Sports adidas group; **Kai Burmeister**, Teamleiter Versorgung-Verträge AOK Nordost

› rer der ubirch GmbH, die neuesten Geräte und deren Funktionsweisen vor. Neben dem praktischen Nutzen der Wearables betonte Noller vor allem auch gesellschaftliche Risiken. So entstünden zum einen dann Probleme, wenn die Anbieter nicht sensibel mit den Daten umgingen und diese an Dritte weitergeben würden. Zum anderen bestehe die Gefahr, dass wichtige Grundsätze wie etwa die Datensparsamkeit zum Beispiel bei Krankenkassen unterlaufen würden. Eine Benachteiligung in Versicherungstarifen auf der Grundlage von Gesundheitsdaten käme einer Endsolidarisierung der Gesellschaft gleich. Daher sei eine umfassende Diskussion über die

Regeln und Werte, die trotz der Digitalisierung erhalten bleiben sollten, besonders wichtig, so der Unternehmer.

Die anschließende Podiumsdiskussion zeigte konträre Meinungen zum Thema Wearables. Burkhard Dümler, Director Development Digital Sports der adidas group, konstatierte einen erhöhten gesellschaftlichen Wunsch nach Selbstvermessung, den die Wirtschaft vor allem durch eine steigende Nachfrage nach Wearables bemerken würde. Wichtig sei den Kunden nicht nur, ihre sportliche Leistung zu kontrollieren, sondern besonders gefragt seien Wearables, die den gesamten Alltag umfassend messbar mach-



Peter Schaar, Vorstand der Europäischen Akademie für Informationsfreiheit und Datenschutz (EAID); **Lena-Sophie Müller**, Geschäftsführerin der Initiative D21 e.V.; **Burkhard Dümler**, Director Development Digital Sports adidas group



Oliver Schenk, Abteilungsleiter G
„Grundsatzfragen der Gesundheitspolitik, Telematik“, Bundesministerium für Gesundheit



Dr. Franz Bartmann, Präsident der Ärztekammer Schleswig-Holstein und Vorsitzender des Ausschusses „Telematik“ der Bundesärztekammer



Stephan Noller,
Gründer und Geschäftsführer der ubirch GmbH

ten. Prof. Dr. Dagmar Borchers, Professorin für Angewandte Philosophie an der Universität Bremen, kritisierte diese Entwicklung und hob hervor, dass die ständige Selbstvermessung die bereits bestehende Leistungsgesellschaft noch weiter verstärken würde. Dem setzte Dr. Bartmann entgegen,

dass das Self-Tracking die Selbstverantwortung und Souveränität des Einzelnen fördern würde. Die subjektive Sicherheit von Patienten, bei denen mithilfe bestimmter Geräte durch kontinuierliche Messung relevanter Parameter eine Verschlechterung im Krankheitsverlauf rechtzeitig erkannt werden kann, würde sich stark erhöhen. Dies verringere außerdem die Frequenz der Arztbesuche, so Bartmann. Die Nutzung von Wearables hätte darüber hinaus einen gesundheitspräventiven Charakter, betonte der Teamleiter Versorgung-Verträge AOK Nordost Kai Burmeister. Dies sei auch der Grund, warum Krankenkassen zunehmend selbst Gesundheits-Apps anbieten und fördern würden. Die Frage, ob Self-Tracking überhaupt gesünder machen würde, brachte Peter Schaar, Vorstand der Europäischen Akademie für Informationsfreiheit und Datenschutz (EAIID), in die Diskussion ein und kritisierte eine dahingehend fehlende Evaluation. Darüber hinaus erhielten die Anbieter man-

cher Wearables schon beim Verkauf einen Freifahrtschein für die Nutzung und Weitergabe der Daten ihrer Käufer. Schaars Forderung nach mehr Datensicherheit in diesem Zusammenhang stimmten die anderen Diskutanten zu.

Die Veranstaltung hat insgesamt gezeigt, dass die Digitalisierung im Bereich Gesundheit große Potenziale bietet. So herrschte weitgehend Einigkeit darüber, dass Big Data in der Medizin große wissenschaftliche Fortschritte, aber auch eine bessere Versorgung des einzelnen Patienten und ein effektiveres Gesundheitssystem hervorbringen könnte. Die private Nutzung von Wearables und Tracking-Methoden wurde unterschiedlich bewertet. Die von Noller und Borchers aufgeworfene Befürchtung, dass der Einsatz von Big Data im Gesundheitsbereich langfristig auch zu einem Gerechtigkeitsproblem führen könnte und daher eine grundsätzliche gesellschaftliche Debatte notwendig sei, fand auch die Zustimmung der anderen Diskutanten. □



Fotos: Tom Maeisa

3.3 Wearables als Tracking-Technologie

Die Gesundheitsindustrie verspricht sich von der Entwicklung und Verbreitung sogenannter Wearables einen großen Impuls für die Nutzbarmachung von Big Data für den Smart-Health-Sektor. Wearables lassen sich allgemein umschreiben als elektronische Geräte, die unmittelbar am Körper getragen werden. Das können Arm- oder Kopfbänder, Armbanduhren, Kopfhörer, Brillen, Kontaktlinsen, auf der Haut angebrachte Pflaster oder auch direkt in Kleidung oder Schuhe eingearbeitete Geräte sein. Durch eingebaute Sensoren und Mikrochips sind sie noch besser als Smartphones in der Lage, die Körperfunktionen des Nutzers zu überwachen und aufzuzeichnen, schon weil sie sich näher am Körper befinden. Die Anbieter gehen davon aus, dass Wearables noch selbstverständlicher vom Nutzer bei sich geführt werden, als das bei Mobiltelefonen der Fall ist. Das gilt insbesondere für Armbänder oder Armbanduhren, die darauf ausgelegt sind, überhaupt nicht abgelegt zu werden, sodass sie theoretisch ein allumfassendes Aktivitätsbild des Anwenders aufzeichnen können. Darüber hinaus sind die Geräte vernetzt, d.h. entweder über die Verbindung mit anderen Geräten wie Mobiltelefonen oder aber selbst mit dem Internet verbunden, um die getrackten Daten auf den Server der Anbieter zu übertragen. Einige Experten haben die Erwartung geäußert, dass die Markteinführung der Apple Watch im Frühjahr 2015 der Technologie zum Durchbruch verhelfen würde, so wie dies durch das iPhone für den Markt der Smartphones geschehen war.⁷¹ Andere Hersteller sind schon länger am Markt vertreten, unter anderen Pebble, Fitbit, Jawbone, Garmin, Samsung, Nike oder LG. Während bislang Fitnessarmbänder dominieren, steht eine Erweiterung der Produktpalette von Wearables in Aussicht. So haben verschiedene Hersteller unter anderem intelli-

gente Funktionsunterwäsche (Heddoko) oder Gürtel (Beltry) angekündigt.

„Die Leute wollen nicht nur eine Messung ihrer sportlichen Leistung – sie wollen sich umfassend selbst beobachten und tracken können. Daher bieten wir zunehmend Wearables an, die das möglich machen.“

Burkhard Dümler, Director
Development Digital Sports adidas group,
öffentliche Veranstaltung Berlin,
30.09.2015

In einer Marktstudie für Deutschland vom März 2015 kommt die Beratungsfirma PricewaterhouseCoopers zu dem Schluss, dass Wearables in den kommenden Jahren dem „Internet der Dinge“ zur Durchsetzung am Markt verhelfen werden, stellt aber fest, dass die Verbreitung in Deutschland bislang noch eher gering ist.⁷² Während im vergangenen Jahr weltweit 150 Millionen Wearables verkauft wurden, hinkt der deutsche Markt mit lediglich 3,6 Millionen Einheiten für ein hoch entwickeltes Land hinterher. In den Vereinigten Staaten hingegen würden bereits heute 21 Prozent der Erwachsenen einen mindestens ein Wearable besitzen. Dennoch sei auch in Deutschland ein großes Interesse an Wearables zu verzeichnen, sodass man davon ausgehen könne, dass die Akzeptanz noch in diesem Jahr signifikant steigen werde. Die Studie prognostiziert, dass Wearables in Deutschland innerhalb der kommenden fünf Jahre im Sinne von Geoffrey Moores Theorie zur Adaption neuer Technologien den Sprung von den „Early Adopters“ hin zur Akzeptanz bei der „Main Majority“ schaffen werden.⁷³

Die Stakeholder

Big Data wird gerade im Hinblick auf den Einsatz im Gesundheitssektor von vielen Experten als „Revoluti-

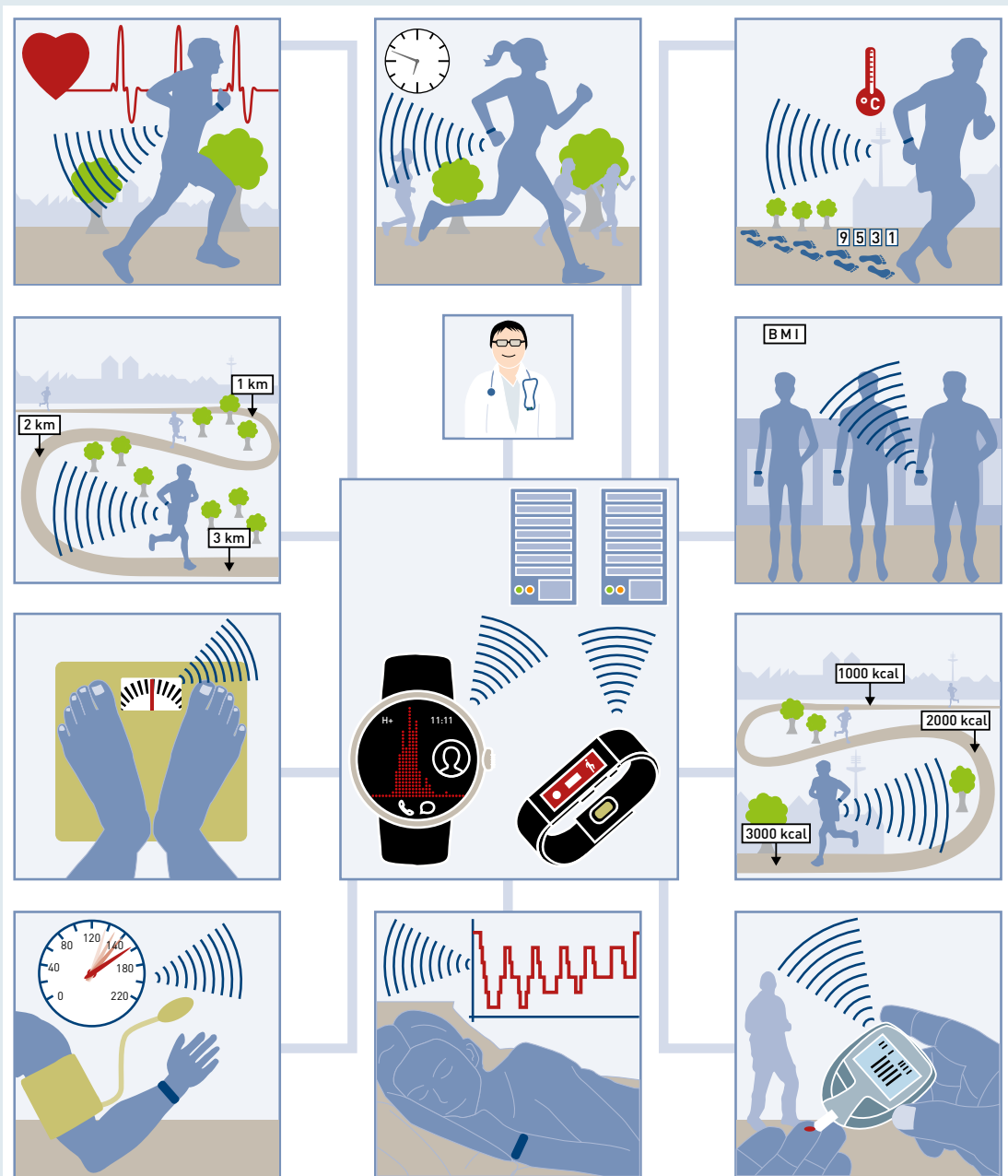
71 Martin Grabmair, „Die Apple Watch verhilft dem Wearable-Markt gerade zum Durchbruch: Was kommt danach? – ein Kommentar“, MacLife Online, 28. April 2015, online: <http://www.maclife.de/news/apple-watch-verhilft-wearable-markt-gerade-zum-durchbruch-was-kommt-jedoch-danach-kommentar-10065179.html>.

72 PWC, „Media Trend Outlook, Wearables: Die tragbare Zukunft kommt näher“, März 2015, online: https://www.pwc.de/de/technologie-medien-und-telekommunikation/assets/pwc-media-trend-outlook_wearables.pdf. Die Studie identifiziert zwei Hauptkategorien von Wearables, zum einen die beschriebenen Geräte zum Körper-Monitoring, zum anderen Human-Computer-Interfaces wie beispielsweise Google Glass. In dem vorliegenden Dokument soll es nur um die erste Kategorie gehen.

73 Ebd., S. 19; siehe dazu Geoffrey A. Moore, „Crossing the Chasm“, HarperCollins, New York 1991.

WEARABLES

Wearables protokollieren Gesundheits- und Aktivitätsdaten des Trägers wie z.B. Anzahl der Schritte, Schlafdauer und Herzfrequenz. Die Daten kann der Anwender einerseits selbst nutzen. Andererseits können sie aber auch für den behandelnden Arzt, die Versicherung oder für die Forschung freigegeben werden.



on“ bezeichnet.⁷⁴ Es ist zu erwarten, dass Wearables bei dieser Entwicklung eine Schlüsselrolle zukommen wird, wenn sich diese erst einmal auf dem Markt etabliert haben: Mehr als jede andere Technologie ermöglichen sie einen direkten und kontinuierlichen Zugang zu den Gesundheitsdaten der Individuen, die dann als Grundlage neuen medizinischen Wissens dienen könnten. Aufgrund dieses Befundes lässt sich eine Vielzahl von Stakeholdern bezüglich der Verbreitung und des Einsatzes von Wearables identifizieren.

- Ärzte gewinnen durch die Nutzung von Wearables die Möglichkeit, den Gesundheitszustand von Patienten besser und präziser zu überwachen, ohne dass es notwendig ist, diese in der Praxis persönlich zu untersuchen. So arbeitet beispielsweise Google an der Entwicklung einer Kontaktlinse, die fortlaufend den Glukosespiegel im Auge von Diabetespatienten messen kann. Die Geräte könnten den behandelnden Arzt automatisiert alarmieren, wenn sich kritische Werte des Patienten so verschlechtern, dass ein unverzügliches Eingreifen nötig wird.⁷⁵ Im Februar 2015 wurde darüber berichtet, dass die Ärzte von mindestens sieben der 17 am besten bewerteten Krankenhäuser in den Vereinigten Staaten begonnen haben, Apples Health-Kit-App zu nutzen, um die relevanten Körperwerte ihrer Patienten zu überwachen.⁷⁶ In diesem Anwendungsbereich geht es nicht um Big Data, sondern unmittelbar um die von dem getragenen Gerät über den Träger gewonnenen Informationen mit dem Zweck, die Diagnose oder Therapie zu verbessern. Big Data kommt erst dann ins Spiel, wenn Ärzte die über Wearables gewonnenen und anschließend aggregierten Daten vieler Patienten mit einer bestimmten Krankheit auswerten und daraus Rückschlüsse ziehen können, um die Behandlung einzelner Patienten zu optimieren – zum Beispiel, welches der verfügbaren Medikamente bei Betroffenen einer bestimmten Altersgruppe am besten oder am verträglichsten wirkt.

„Bestimmte Wearables erhöhen die Sicherheit und Autarkie vor allem chronisch kranker Patienten, da Verschlechterungen im Verlauf bei kontinuierlicher Messung entscheidender Parameter, z.B. drohende Blutzuckerentgleisungen oder Hochdruckkrisen, rechtzeitig und nicht mehr oder weniger zufällig anlässlich eines vereinbarten Arzttermins erkannt werden.“

Dr. Franz Bartmann, Präsident der Ärztekammer Schleswig-Holstein und Vorsitzender des Ausschusses „Telematik“ der Bundesärztekammer, öffentliche Veranstaltung Berlin, 30.09.2015

- Krankenhäuser können allgemein von den Möglichkeiten des Einsatzes von Big Data profitieren. Aus großen Datensätzen von Patienten lassen sich Erkenntnisse gewinnen, um die verfügbaren Ressourcen besser zu verteilen. Diese Patientendaten lassen sich unter anderem aus dem Einsatz von Wearables gewinnen. So ist zumindest theoretisch denkbar, dass eine Auswertung der Daten aller Nutzer im Einzugsbereich eines Krankenhauses Rückschlüsse darauf zulässt, mit welcher Art von Krankheiten verstärkt zu rechnen ist – zum Beispiel, weil aufgrund des hohen Durchschnittsalters der Bevölkerung in der Region vermehrt Herzrhythmusstörungen verzeichnet werden.
- Ähnliches lässt sich für Gesundheitsbehörden prognostizieren. Allerdings ist hier ein unmittelbarer Nutzen davon abhängig, dass Wearables in den kommenden Jahren tatsächlich eine große Verbreitung finden. Erst wenn eine ausreichend große Anzahl an Bürgern im Zuständigkeitsbereich der jeweiligen Behörde am Tracking der eigenen Gesundheitsdaten teilnimmt, können die dann entstehenden Datenmengen dazu genutzt werden, für bestimmte Krankheiten Frühwarnsysteme einzurichten.

⁷⁴ Vgl. z.B. Viktor Mayer-Schönberger/Kenneth Cukier, „Big Data. Die Revolution, die unser Leben verändern wird“, Redline, München 2013.

⁷⁵ Nanette Byrnes, „Mobile Health’s Growing Pains“, MIT Technology Review Online, 21. Juli 2014, online: <http://www.technologyreview.com/news/529031/mobile-healths-growing-pains/>.

⁷⁶ Tobias Költzsch, „Immer mehr Krankenhäuser nutzen Apples Gesundheits-App“, Golem.de, 5. Februar 2015, online: <http://www.golem.de/news/healthkit-immer-mehr-krankenhaeuser-nutzen-apples-gesundheits-app-1502-112175.html>.

„Privatsphäre ist mir als Privatperson wichtig. Insbesondere bei Gesundheitsdaten ist dieser Schutz sehr ernst zu nehmen.“

Peter Ohnemus, Präsident, Gründer und CEO dacadoo AG, Konsultation]

- Wie bereits angedeutet, verspricht sich die medizinische Forschung großen Nutzen vom Einsatz des Trackings von Gesundheitsdaten. War es bislang mit sehr viel Aufwand verbunden, Probanden für medizinische Studien zu gewinnen, weil die persönliche Teilnahme in einer Klinik eine hohe Schwelle für Freiwillige bedeutet⁷⁷, können die notwendigen Datensätze nun sehr leicht und durch die Nutzer selbst gewonnen werden, indem entsprechende Apps genutzt werden. Ist diesbezüglich schon jetzt eine beachtliche Bereitschaft seitens der Smartphone-Besitzer zu beobachten, wie die große Teilnahme an den ersten zum Download bereitgestellten Apps auf der Basis von Apples ResearchKit zeigte⁷⁸, kann realistisch davon ausgegangen werden, dass die weitere Verbreitung von Wearables dazu führen wird, dass diese Daten von mehr Menschen bereitwillig der Forschung zur Verfügung gestellt werden. Die damit verbundenen Chancen werden im nächsten Abschnitt näher ausgeführt.

„Ich glaube, dass es für die Nutzung von Daten viele interessante Ansätze gibt, und diese sollten vor allem interdisziplinär weiterverfolgt werden. Zentral ist es nicht nur zu fragen, was technisch realisierbar ist, sondern was die Zielsetzung ist.“

Dr. Silke Jandt, Vertreterin des Lehrstuhls für Öffentliches Recht, Informationstechnologierecht und Rechtsinformatik an der Universität Passau, Interview

- Krankenversicherungen können die mittels Wearables gewonnenen Informationen über den Gesundheitszustand der Versicherungsnehmer nutzen, um ihre Preismodelle zu optimieren und somit Kosten zu sparen. Die Analyse großer Datensätze ermöglicht es, Kostenrisiken besser abzuschätzen, indem die Einzeldaten miteinander kombiniert werden, wodurch ein genaueres Bild über den Gesundheitszustand und die -prognose individueller Versicherungsnehmer erstellt werden kann. Einige Versicherungen haben begonnen, ihren Kunden Bonusleistungen anzubieten, wenn diese ihre Gesundheitsdaten aufzeichnen und übermitteln. Dadurch sollen sie auch motiviert werden, durch Fitness und bessere Ernährung mehr auf ihre Gesundheit zu achten. Ein besonders signifikantes Beispiel stellt in dieser Hinsicht das gemeinsame Geschäftsmodell der australischen Supermarktkette Coles mit der Krankenversicherung Medibank dar. Das Lebensmittelunternehmen hat ein Kundenbindungsprogramm, das nicht nur das Kaufverhalten der teilnehmenden Kunden registriert, sondern mittels Wearables auch deren Gesundheits- und Fitnessdaten. Das Unternehmen kooperiert mit der Versicherung Medibank, die den Kunden über ein Bonusprogramm Anreize gibt, regelmäßig ihre Daten auf eine Online-Plattform hochzuladen. Mit diesem Modell können die Gesundheitsdaten mit Daten zum Einkaufsverhalten zum Beispiel von Alkohol, Zigaretten oder fettigen und kalorienreichen Lebensmitteln zusammengeführt werden, um aus den kombinierten Daten individuelle Gesundheitsprognosen zu erstellen und dann die angebotenen Versicherungsmodelle entsprechend anzupassen.⁷⁹ Die gesellschaftlichen Risiken, die mit einer solchen Praxis verbunden sind, werden im nächsten Abschnitt thematisiert.

⁷⁷ Vgl. Natasha Lomas, „ResearchKit. An ‚Enormous Opportunity‘ for Science, Says Breast Cancer Charity“, TechCrunch, 14. März 2015, online: <http://techcrunch.com/2015/03/14/researchkit-share-the-journey/>.

⁷⁸ Zen Chu, „Why Apple’s ResearchKit signals a Golden Age for health care“, Fortune Online, 27. März 2015, online: <http://fortune.com/2015/03/27/why-apples-researchkit-signals-a-golden-age-for-health-care/>.

⁷⁹ Vgl. <http://www.coles.com.au/about-coles/news/2013/11/25/coles-and-medibank-reward-customers>.

„Predictive Analysis wird es den Krankenkassen ermöglichen, individualisierte Risikoprofile zu erstellen und mit sehr granularen Individualтарifen zu hinterlegen. Man darf davon ausgehen, dass die Kassen versuchen werden, diese technologischen Möglichkeiten auch zu nutzen.“

Dr. Bernhard Rohleder,
Hauptgeschäftsführer Bitkom e.V.,
2. Expertenworkshop, 15.06.2015

- Die Plattformanbieter werden als Aggregatoren von Einzeldaten tätig. Sie können mit den daraus gewonnenen Ergebnissen handeln oder darauf aufbauend weitere Produkte anbieten. So haben bereits seit mehreren Jahren viele größere Softwareentwicklungs-Unternehmen wie IBM, SAP oder Oracle begonnen, nicht nur die Speicherung großer Datenmengen in ihren Serverfarmen anzubieten, sondern als eigentliche Serviceleistung die Aggregation und Auswertung der Daten zum Geschäftsmodell gemacht. Kunden können dann unmittelbar die Rechenleistung der Big-Data-Plattformen nutzen und die entsprechenden Anwendungen „in die Cloud“ auslagern – im Gesundheitsbereich bietet IBM mit seiner Health-Sparte einen solchen Dienst an, bei dem das Watson-System für die Verarbeitung der eingespeisten Daten zum Einsatz kommt.⁸⁰ In jüngerer Zeit haben die Unternehmen in diesem Bereich ihren Umsatz signifikant steigern können.⁸¹

„Manche Hersteller von Wearables nötigen den Kunden schon beim Kauf die Zusage ab, ihre echten Daten anzugeben und einzuwilligen, dass das Unternehmen die Daten nutzen und weitergeben darf.“

Peter Schaar, Vorstand der Europäischen Akademie für Informationsfreiheit und Datenschutz (EAIID), öffentliche Veranstaltung Berlin, 30.09.2015

- Schließlich sind natürlich auch die Nutzer selbst zu den Stakeholdern im Bereich Wearables zu zählen. Zentral erscheint hierbei, dass beachtliche kulturelle Unterschiede zu beobachten sind, in welcher Form Menschen bereit sind, ihre mittels Wearables generierten Daten freizugeben. So berichtet die erwähnte Studie von Pricewaterhouse-Coopers, dass 70 Prozent der US-amerikanischen Erwachsenen in einer Befragung im Herbst 2014 angaben, sie würden ein Wearable zur Überwachung ihrer Gesundheitsdaten tragen, wenn dies zur Senkung ihrer Versicherungsprämie führen würde. In Deutschland hingegen könnte nur ein Drittel der Befragten durch finanzielle Anreize dazu motiviert werden, persönliche Daten freizugeben. Nur 26 Prozent würden ihrer Krankenkasse die aufgezeichneten Daten anvertrauen, aber immerhin 63 Prozent ihrem Hausarzt oder einem Krankenhaus. Datenschutz und -sicherheit spielen für deutsche Nutzer eine große Rolle. So gaben 98 Prozent an, dass ihnen dieser Aspekt besonders wichtig sei und einen entscheidenden Faktor für die Kaufentscheidung darstelle.⁸² Insgesamt dienen Wearables in Deutschland bislang vor allem dazu, die eigenen Gesundheitsdaten für rein private Zwecke zu überwachen (sogenanntes Self-Monitoring). Zwei Drittel der hiesigen Nutzer verbinden mit den Geräten die Erwartung, dass sie bei der individuellen Gesundheitsvorsorge helfen. Da hierfür nicht zwingend Daten geteilt werden müssen, koppeln nur knapp 60 Prozent der Anwender ihr Wearable überhaupt mit weiteren Geräten, um Datenaustausch zu ermöglichen.⁸³

Die Chancen von Tracking mittels Wearables

Die Interessen der Stakeholder beim Einsatz von Wearables zur Gewinnung von Gesundheitsdaten weisen bereits auf die Chancen hin, die mit der Technologie verbunden werden.

Mit der Nutzung von Wearables sind Erwartungen verknüpft, dass durch das Monitoring und Tracking der

⁸⁰ <http://www.ibm.com/smarterplanet/us/en/ibmwatson/health/>.

⁸¹ Vgl. <http://de.statista.com/statistik/daten/studie/302303/umfrage/umsatz-der-fuehrenden-anbieter-von-big-data-loesungen-weltweit/>.

⁸² PWC, „Wearables“, s.o.

⁸³ Ebd., S. 8.

Gesundheitsdaten Diagnosen präziser ausfallen und Therapien besser an die individuellen Bedürfnisse des einzelnen Patienten angepasst werden können. Je mehr Informationen über die behandlungsbedürftige Person bekannt sind, so die zugrunde liegende These, desto höher die Heilungschancen. Zu diesem Zweck können sowohl Small Data, also die Körperwerte des Patienten selbst, als auch aus Big Data gewonnene Erkenntnisse fruchtbar gemacht werden: Gerade die Aggregation und Auswertung von Daten vieler Personen soll dazu führen, dass mehr über die Krankheit und mögliche Therapieformen bekannt wird. Genauso können auf diese Weise Vorsorge und Früherkennung von Krankheiten verbessert werden. Hinzu kommt der Aspekt der Fitness. Wearables arbeiten auch mit sogenannten Gamification-Elementen⁸⁴, die Nutzer dazu motivieren sollen, gesünder zu leben. Zu nennen wären beispielsweise von den Apps vorformulierte „Ziele“, die erreicht werden sollen und an die in regelmäßigen Abständen erinnert wird.

„Durch bestimmte Muster in Big-Data-Analysen können Vorhersagen getroffen werden, die Personen mit bestimmten Gesundheitsrisiken konfrontieren, die aber gegebenenfalls gar nicht eintreten. Wir sollten uns rechtzeitig Gedanken machen, wie wir damit umgehen wollen und wie weit Big Data im Gesundheitswesen gehen soll.“

Dr. Franz Bartmann, Präsident der Ärztekammer Schleswig-Holstein und Vorsitzender des Ausschusses „Telematik“ der Bundesärztekammer, öffentliche Veranstaltung Berlin, 30.09.2015

Sollte sich die Erwartung bezüglich verbesserter Diagnosen und Therapien erfüllen, so würde der mittels Wearables erstellte Datenbestand die Gesundheit der Gesamtbevölkerung steigern helfen, wodurch wiederum das Gesundheitssystem insgesamt entlastet würde. Der gleiche Effekt kann sich auch durch

den Fitness-Aspekt ergeben. Die spielerischen Elemente der mit den Wearables verknüpften Apps könnten dazu führen, dass individuell mehr auf die Gesundheit geachtet wird. Auch dies kann in der Summe in einer Verbesserung der gesellschaftlichen Gesundheit resultieren. Dementsprechend erhoffen sich auch die Krankenversicherungen mittel- bis langfristig eine Entlastung ihrer Budgets, wenn die Preismodelle besser an die tatsächlichen versicherungsrelevanten Umstände der einzelnen Versicherungsnehmer angepasst werden können.⁸⁵ Schließlich könnte auch die Pharmaindustrie profitieren, wenn die aggregierten Daten für die Entwicklung neuer oder die Optimierung bereits auf dem Markt vorhandener Medikamente zur Verfügung gestellt werden.

„Smart-Health-Produkte können auch das individuelle Verantwortungsbewusstsein erhöhen.“

Dr. med. Peter Langkafel, Gründer und Geschäftsführer der Healthcubator GmbH, 2. Expertenworkshop, 15.06.2015

Mit letzterem Aspekt eng verknüpft sind die Erwartungen der medizinischen Forschungsgemeinschaft an die Entwicklungen von Big Data. Bislang war es für Universitätskliniken und Forschungsinstitute ein großes Problem, eine ausreichende Anzahl an Probanden für Studien zu akquirieren. Teilnehmer mussten mühsam mit Flugblättern oder Anzeigen rekrutiert werden, aber selbst bei größtem Aufwand blieb die Datenbasis am Ende oft schmal. Gerade in der Medizin ist eine breite statistische Basis aber essenziell. Je mehr Personen ihre Daten zur Verfügung stellen, desto akkurater lässt sich die Gesamtbevölkerung abbilden, wodurch die Studienergebnisse an Aussagekraft gewinnen. Die Vorstellung des ResearchKit von Apple zusammen mit der Apple Watch schickt sich an, dieses Problem zu lösen.⁸⁶ Auf Basis einer Open-Source-Plattform können Forschungseinrichtungen Anwendungen für das iPho-

⁸⁴ Gamification: die Anwendung spieltypischer Elemente und Prozesse in spielfremdem Kontext. Zu diesen spieltypischen Elementen gehören unter anderem Erfahrungspunkte, Highscores, Fortschrittsbalken, Ranglisten, virtuelle Güter oder Auszeichnungen (Quelle: <http://de.wikipedia.org/wiki/Gamification>).

⁸⁵ Dass Big-Data-gestützte Preisdiskriminierung bzw. -individualisierung keineswegs auf Versicherungen beschränkt sein wird, sondern sich bis hin zu Lebensmitteln im Supermarkt erstreckt, zeigt anschaulich Hannes Grassegger, „Jeder hat seinen Preis“, Zeit Online, 27. Oktober 2014, online: <http://www.zeit.de/wirtschaft/2014-10/absolute-preisdiskriminierung/komplettansicht>.

⁸⁶ Vgl. <https://www.apple.com/researchkit/>.

ne und die Apple Watch entwickeln, mittels derer Nutzer die für das jeweilige Projekt benötigten Daten übermitteln können. Es können Apps programmiert werden, die je nach Bedarf Zugriff auf Beschleunigungssensor, Gyroskop, GPS-Modul und das Mikrofon des Mobiltelefons bekommen. Die ersten fünf Anwendungen, die zusammen mit dem Start der Plattform veröffentlicht wurden, lassen Nutzer an Testreihen zu Parkinson, Diabetes, Herz-Kreislauf-Krankheiten, Asthma und Brustkrebs teilnehmen. Die Methode der Datengewinnung mittels der Apple Watch, also einem Wearable, minimiert die Hürden der Teilnahme an einer medizinischen Studie. Es ist nicht länger notwendig, in ein Universitätskrankenhaus zu gehen und mit großem Zeitaufwand Formulare auszufüllen. Wenn das Gerät ständig getragen wird und während der Erledigung von alltäglichen Aktivitäten fortlaufend Gesundheitsdaten getrackt werden können, entsteht eine Menge an relevanten Daten, die durch traditionelle Forschungsmethoden gar nicht gewonnen werden könnte.

„Big Data ist ein omnipräsentes Phänomen. Es gibt riesige Datenschätze, aber die Zugänge dazu bleiben verwehrt.“

Prof. Dr.-Ing. Ina Schieferdecker,
Leiterin des Fraunhofer-Instituts für
Offene Kommunikationssysteme
(Fraunhofer FOKUS), Konsultation

„Ob uns Wearables wirklich gesünder machen, ist nicht eindeutig belegt. Hier brauchen wir zunächst umfassende Evaluationen.“

Peter Schaar, Vorstand der Europäischen
Akademie für Informationsfreiheit und
Datenschutz (EAID), öffentliche
Veranstaltung, 30.09.2015

Die Gesamtschau all dieser Aspekte rechtfertigt es in der Tat, von einer Revolution des Gesundheitssystems durch Big Data zu sprechen. Ihren folgenreichsten Ausdruck erfährt diese Umwälzung in der grundlegenden Transformation der Rolle des Individuums. Wenn diesem anhand von Wearables die Möglichkeit gegeben wird,

stets über den eigenen Gesundheitszustand informiert zu sein, ist der Einzelne nicht mehr länger nur passiver Empfänger von Gesundheitsversorgung, sondern wird zum aktiven Gestalter – jedenfalls potenziell. Krankheiten sind unter dieser Prämisse weniger ein hinzunehmender Schicksalsschlag als ein in der Verantwortung des Betroffenen liegender Umstand, den es mittels Selbstüberwachung zu verhindern gilt. Das gilt natürlich nicht für alle Krankheiten. Doch gerade viele der heutigen „Volkskrankheiten“ werden mit dem Lebensstil moderner, postindustrieller Gesellschaften in Verbindung gebracht. Wenn es jeder Person durch das Tragen von Wearables künftig möglich ist, Warnzeichen frühzeitig zu erkennen und anschließend Gewohnheiten beispielsweise hinsichtlich Ernährung und Fitness rechtzeitig zu ändern, dann ist davon auszugehen, dass sich ein entsprechender gesellschaftlicher Imperativ langfristig herausbilden wird. Das „Quantified Self“ – also die Vermessung des Menschen mittels Apps und Wearables⁸⁷ – würde so zur Norm.

„Aus Sicht einer Krankenkasse sind neben dem Datenschutz die Akzeptanz und der Nutzen für die Kunden das Entscheidende.“

Dr. Kai Kolpatzik, Leiter Prävention
beim AOK-Bundesverband,
2. Expertenworkshop, 15.06.2015

Die Risiken von Tracking mittels Wearables

Dass eine solche Entwicklung nicht nur vorteilhaft ist, sondern auch problematische Konstellationen mit sich bringt, liegt auf der Hand. Personen, die ihre Gesundheitsdaten mittels Wearables tracken, gewinnen an Autonomie hinzu, weil sie die Kontrolle über ihr eigenes Wohlbefinden und damit Verantwortung für die eigene Gesundheit übernehmen. Zugleich jedoch vergrößert sich die Gefahr der Fremdbestimmtheit, wenn durch die konstante Überprüfbarkeit und Vergleichbarkeit der Daten gesellschaftlich neu definiert wird, welche körperlichen Zustände als gesund, also „richtig“, und welche als krank, also „falsch“, gelten. Die auf den ersten Blick rein deskriptiven Daten werden so zu Kennzahlen für akzeptier-

⁸⁷ Vgl. Alina Schadwinkel, „Die 10.000 Fragezeichen“, Zeit Online, 20. April 2015, online: <http://www.zeit.de/wissen/gesundheit/2015-04/quantified-self-fitness-gesundheit-wissenschaft>.

tes Verhalten.⁸⁸ Die auf diese Weise neu entstandenen Normen erzeugen einen Anpassungsdruck, der Abweichung sanktioniert und so zur Befolgung zwingt.⁸⁹ Im Extremfall könnte daraus die Aufkündigung des Solidarprinzips bei der Krankenversicherung folgen. Die ökonomischen Konsequenzen von Krankheit trägt eine Gesellschaft unter anderem deshalb gemeinschaftlich, weil potenziell jedes Mitglied betroffen sein könnte. Wer aber krank wird, obwohl Warnzeichen verfügbar gewesen wären, der kann unter diesen Vorzeichen eben nicht mehr mit dem gleichen Argument auf die Solidargemeinschaft zählen. Das kann zum Beispiel dazu führen, dass bestimmte, als ungesund definierte Verhaltensweisen künftig dadurch „bestraft“ werden, dass die Nutzer höhere Versicherungsbeiträge zahlen müssen, wenn sie mittels Trackings „überführt“ werden, einen ungesunden Lebensstil zu führen. Eine solche unmittelbare Verknüpfung von Verhalten (ungesunder Lebensstil) und Konsequenz (höhere Beiträge) ignoriert aber, dass viele der Gesundheit abträgliche Verhaltensweisen beispielsweise auch gesellschaftliche Ursachen haben können, wie zum Beispiel ungleich verteilten Zugang zu Bildung oder nicht ausreichende finanzielle Budgets für teurere gesunde Nahrungsmittel bei sozial schwachen Schichten.⁹⁰ Auf diese Weise könnten bestimmte Bevölkerungsteile durch die mittels Tracking neu entstandenen gesellschaftlichen Normen in Bezug auf Gesundheit diskriminiert werden.

„Neue Systeme im Gesundheitsbereich sind denkbar und können Vorteile haben. Es ist durchaus fair, wenn Anreize geschaffen werden, sich gesundheitsfördernder zu verhalten und dafür einen geringeren Versicherungsbeitrag zu leisten. Gleichzeitig sollte das Solidarprinzip erhalten bleiben, bei dem benachteiligte Personen nicht abgestraft werden und in gleichem Umfang Leistungen erhalten.“

Florian Schumacher, Gründer Quantified Self Deutschland und Digital Health Consultant iic-solutions, Konsultation

„Mit dem Einsatz von Big Data in allen Lebensbereichen entsteht die Gefahr einer Entsolidarisierung. Daher brauchen wir nicht nur eine technische, sondern auch eine gesellschaftliche Diskussion über unsere Werte und Regeln.“

Stephan Noller, Gründer und Geschäftsführer der ubirch GmbH, öffentliche Veranstaltung Berlin, 30.09.2015

An diese Überlegung schließt sich unmittelbar das Problem der Freiwilligkeit an. Sollten Wearables, die Gesundheitsdaten tracken, in naher Zukunft genauso weit verbreitet sein wie Smartphones heute, dann muss davon ausgegangen werden, dass beispielsweise durch eine veränderte Preispolitik der Versicherungen der Druck auf die individuellen Nutzer steigen wird, die generierten Daten zu übermitteln, um sie gesamtgesellschaftlichen Zwecken zur Verfügung zu stellen. Auch wenn die meisten Hersteller versichern, die von den Nutzern erzeugten Daten nicht zu verwerten, verschwimmen bei Wearables schon heute die Grenzen zwischen Small und Big Data zunehmend. Die Daten, die beim Anwender zur Überprüfung der eigenen Fitness oder des Gesundheitszustandes entstehen, werden vermehrt von den Anbietern von Wearables als potenzielle Vermögenswerte betrachtet, die sich zu Geld machen lassen, wenn sie mit den Daten anderer Nutzer zusammengeführt werden, um Big-Data-Anwendungen zu ermöglichen.⁹¹ Sobald dies geschieht, werden die privaten, mitunter als intim einzustufenden Daten für Zwecke genutzt, die über die Interessen der einzelnen Person hinausgehen. Ein Beispiel hierfür ist die Webseite Patients Like Me.⁹² Nutzer können sich auf dieser Plattform mit anderen Betroffenen über ihre Krankheiten austauschen und werden dazu motiviert, ihre insoweit relevanten Gesundheitsdaten hochzuladen. Auf diese Daten können der Anbieter, aber auch Dritte wie Pharmafirmen zugreifen,

⁸⁸ Vgl. Stefan Selke, „Lifelogging. Wie die digitale Selbstvermessung unsere Gesellschaft verändert“, Berlin 2014.

⁸⁹ Vgl. Anita L. Allen, „Dredging up the Past: Lifelogging, Memory, and Surveillance“, The University of Chicago Law Review, Vol. 75, 2008, S. 47.

⁹⁰ Evgeny Morozov, „The rise of data and the death of politics“, The Guardian Online, 20. Juli 2014, online: <http://www.theguardian.com/technology/2014/jul/20/rise-of-data-death-of-politics-evgeny-morozov-algorithmic-regulation>.

⁹¹ Deborah Lupton, „Self-Tracking Modes: Reflexive Self-Monitoring and Data Practices“, Paper for „Imminent Citizenships: Personhood and Identity Politics in the Informatic Age“, Canberra 2014, S. 5, online: <http://ssrn.com/abstract=2483549>.

⁹² <https://www.patientslikeme.com/>.

um aus den aggregierten Datensätzen Erkenntnisse für die eigenen Geschäftsmodelle zu gewinnen.⁹³ Natürlich können auch die Nutzer selbst die durch Activity Tracking gewonnenen Informationen ausdrücklich zu dem Zweck zur Verfügung stellen, um die Verwertung durch Big-Data-Anwendungen zu ermöglichen. Dies geschah beispielsweise massenhaft nach der Einführung der Open-Software-Umgebung ResearchKit durch Apple. Zum Start der Plattform wurde unter anderem die App „MyHeart Counts“ vorgestellt, die von der Stanford University entwickelt worden war. Innerhalb weniger Tage wurde die App in den Vereinigten Staaten und Kanada fast 53.000-mal heruntergeladen. Anschließend willigten 22.000 Nutzer ein, an der Studie teilnehmen zu wollen, obwohl sie als Gegenleistung lediglich eine Analyse ihrer Gesundheitsdaten versprochen bekamen.⁹⁴ Dagegen lassen sich zunehmend auch Fälle beobachten, in denen Nutzer mehr oder minder stark genötigt werden, ihre Gesundheitsdaten zu übermitteln.⁹⁵

„Gesundheitstracking könnte das Ende der Solidargemeinschaft bedeuten. Wenn alles vom Individuum abhängt, wird das Risiko nicht mehr auf mehrere Schultern verteilt.“

Prof. Dr. Tobias O. Keber, Professor für Medienrecht und Medienpolitik, Institut für Digitale Ethik (IDE) der Hochschule der Medien Stuttgart, Interview

„Was haben wir für ein Bild von Gesundheit und Krankheit, wenn wir es nur auf Zahlen reduzieren? Gesundheit ist auch geprägt durch Biografie, Umfeld und persönliche Erlebnisse – was in der Statistik keine Rolle spielt.“

Prof. Dr. Dagmar Borchers, Professorin für Angewandte Philosophie an der Universität Bremen, öffentliche Veranstaltung, 30.09.2015

Modi des Self-Trackings

Ausgehend von einer kritischen Betrachtung bezüglich der Frage der Freiwilligkeit von Self-Tracking, hat die australische Soziologin Deborah Lupton insgesamt fünf verschiedene Modi identifiziert, um Anwendungsfälle des Trackings in Bezug auf die Weitergabe sensibler Daten zu klassifizieren: privat, gemeinschaftlich, angestoßen, aufgezwungen und ausgenutzt. Dabei handelt es sich nicht um klar voneinander abgegrenzte oder abgrenzbare Kategorien. Die Einordnungen überlappen und kreuzen sich vielmehr.⁹⁶

Privates Tracking bedeutet, dass die von Wearables getrackten Daten nur für die eigenen Zwecke des Nutzers aufgezeichnet werden. Sie werden entweder gar nicht oder nur mit einem klar abgegrenzten Kreis an Personen geteilt, also beispielsweise mit Familienmitgliedern oder Freunden. Nach dem Willen des Anwenders verbleiben die Daten also bei ihm. Sie sollen demnach gerade nicht zu Big Data aggregiert werden. Bei diesem Modus des Trackings, das auch als „Self-Tracking“ beschrieben werden kann, geht es dem Nutzer darum, systematisch Informationen zur eigenen Ernährung, allgemeinen Gesundheit und Aktivität zu gewinnen, um daraus Erkenntnisse für die Verbesserung dieser Werte und des eigenen Wohlbefindens zu gewinnen. Die Möglichkeit, durch das Tracking solcher Werte dazu motiviert zu werden, mehr auf die eigene Gesundheit zu achten und die eigene Lebensqualität zu steigern, ist die klassische, ursprüngliche Form der Nutzung solcher Geräte und Softwareanwendungen und hat überhaupt erst zu der Gründung von Bewegungen wie „Quantified Self“ geführt. Auswirkungen für die Gesellschaft als Ganze sind in diesem Sinne immer nur mittelbare und beiläufige Folgen des Trackings, nicht dessen Zweck. Die Vermessung des eigenen Körpers soll dem Einzelnen die Möglichkeit geben, hochgradig persönliche Zusammenhänge zu erkennen und damit verbundene gesundheitliche Fragestellungen ganz individuell und autonom zu beantworten.

⁹³ Lupton, „Self-Tracking Modes“, s.o., S. 11.

⁹⁴ Lomas, „ResearchKit“, s.o.

⁹⁵ Das geschieht vor allem in den Vereinigten Staaten vermehrt durch Arbeitgeber; dazu genauer sogleich.

⁹⁶ Lupton, „Self-Tracking Modes“, s.o., S. 5 ff.

Eine Erweiterung des privaten Trackings ist das gemeinschaftliche. Hier bleibt der Zweck gleich, es geht in erster Linie weiter um die Verbesserung der persönlichen Gesundheit. Aus verschiedenen Motivationen heraus können Nutzer sich dafür entscheiden, ihre Daten mit anderen in einer Community zu teilen. Das kann aus einem Wettbewerbsgedanken heraus geschehen, aber auch, um die eigenen Werte mit denen anderer Self-Tracker zu vergleichen. Wenn sich auf diese Weise Standards oder Richtwerte herausbilden, so sind diese im Normalfall spezifisch für die jeweilige Gemeinschaft. Auch für das gemeinschaftliche Tracking kann die „Quantified Self“-Bewegung als Grundmodell dienen. Schon kurz nach der Gründung in den Vereinigten Staaten im Jahr 2007 bildeten sich Gruppen in den deutschen Metropolen. Inzwischen gibt es in fast jeder größeren deutschen Stadt eine „Quantified Self“-Community, die sich dem regelmäßigen Austausch unter den Mitgliedern über Methoden und Technologien, aber eben auch über die gewonnenen Daten selbst und ihre Interpretation und Auswertung verschrieben hat.

Als dritte Kategorie identifiziert Lupton das angestoßene Tracking. Dieses umschreibt das Aufzeichnen der eigenen Vitalitätsdaten, das nicht durch den Nutzer selbst motiviert ist, sondern durch eine andere Person oder Institution. Das kann zum Beispiel der Arbeitgeber des Nutzers sein, der seine Mitarbeiter dazu anhalten will, mehr auf ihre Gesundheit zu achten, um Krankheitstage zu verringern und damit die Produktivität der Belegschaft zu erhöhen. Dieses Modell greift insbesondere in den Vereinigten Staaten immer mehr um sich, und die Anbieter von Wearables haben entsprechend begonnen, auf diese Entwicklung zu reagieren. So hat beispielsweise der Hersteller von Armbändern Fitbit ein Geschäftsmodell entwickelt, das es Arbeitgebern ermöglicht, mit dem Anbieter spezielle Tarife auszuhandeln, damit sie Fitbit-Geräte für firmeninterne Fitnessprogramme verwenden.

„Krankenkassen bieten bereits selbst Apps für den Gesundheitsbereich an, mit denen vor allem die Prävention angesprochen wird. Wir erreichen damit Menschen, die bisher durch analoge Maßnahmen nicht ausreichend erreicht werden konnten.“

Kai Burmeister, Teamleiter
Versorgung-Verträge AOK Nordost,
öffentliche Veranstaltung Berlin,
30.09.2015

Aufgezwungenes Tracking bildet die nächste Stufe der Aufzeichnung von Gesundheitsdaten. Im Gegensatz zum angestoßenen Tracking fehlt hier das Moment der Freiwilligkeit ganz. Denkbar ist, dass dies auf staatliche Anordnung geschieht. So könnten theoretisch verurteilte Straftäter, die auf Bewährung frei sind, zum Tragen von Wearables gezwungen werden, damit auf diese Weise ihr Drogen- oder Alkoholkonsum überwacht werden kann – auch wenn es sich bei diesem Szenario bislang noch um Fiktion handelt. Genauso gut könnten Nutzer aber auch von privater Seite zur Verwendung von Fitnessarmbändern oder ähnlichen Geräten genötigt werden. Ein Arbeitgeber beispielsweise könnte es zur Pflicht machen, dass Arbeitnehmer an internen Programmen zur Erhöhung der Fitness teilnehmen müssen, und sie zwingen, Fitnessarmbänder zu tragen. Ob eine solche Praxis in Deutschland arbeitsrechtlich zulässig wäre, ist fraglich. In den Vereinigten Staaten hingegen wird es bereits heute vielen Arbeitnehmern zunehmend schwer gemacht, sich solchen Überwachungsmaßnahmen zu verweigern.⁹⁷

Ausgenutztes Tracking schließlich beschreibt den Fall, dass durch Self-Tracking gewonnene Daten – egal, ob diese freiwillig oder aufgezwungen generiert wurden – von Dritten wirtschaftlich verwertet werden. Die persönlichen Daten Einzelner werden so zu Wirtschaftsgütern.⁹⁸

⁹⁷ Vgl. Suzanne McGee, „How employers tracking your health can cross the line and become Big Brother“, The Guardian Online, 1. Mai 2015, online: <http://www.theguardian.com/lifeandstyle/us-money-blog/2015/may/01/employers-tracking-health-fitbit-apple-watch-big-brother>; Alex Rosenblat u.a., „Workplace Surveillance“, Data & Society Research Institute, 2014, online: <http://www.datasociety.net/pubs/fow/WorkplaceSurveillance.pdf>.

⁹⁸ Siehe oben das Beispiel der Webseite von Patients Like Me.

Während die ersten zwei Modi des Activity Tracking als grundsätzlich unproblematisch einzustufen sind, insofern, als sie das Prinzip der Freiwilligkeit nicht verletzen, muss schon bei der angestoßenen Weitergabe von Gesundheitsdaten genau geklärt werden, welche Anforderungen an die Zustimmung des Nutzers zu stellen sind. Gesundheitsdaten werden schon deshalb als besonders sensibel angesehen, weil viele Krankheiten (wie zum Beispiel HIV oder Depressionen) gesellschaftlich stigmatisiert sind.⁹⁹ Entsprechend geben schon heute 62 Prozent der deutschen Nutzer von Wearables an, diese würden zu sehr in die Privatsphäre eindringen.¹⁰⁰

Dieses Problem dürfte sich in Zukunft noch verschärfen, da erwartet wird, dass die Geräte und Anwendungen noch invasiver werden. So hat das Unternehmen Ginger.io beispielsweise eine App entwickelt, die es anhand der Bewegungsdaten und Sprachaufzeichnungen bei manisch-depressiven Patienten ermöglichen soll, manische Episoden frühzeitig zu erkennen.¹⁰¹

Die Einwilligung zur Weitergabe und Nutzung von Daten ist häufig bewusst in den allgemeinen Nutzungsbedingungen versteckt. In der Folge ist es vielen Nutzern also nicht klar, dass sie bei Verwendung von Tracking-Geräten oder -Software einer solchen weiter gehenden Verwertung ihrer Daten überhaupt zugestimmt haben.

Das Problem der Freiwilligkeit

Neben der Frage der informierten Zustimmung ist die der Freiwilligkeit der Datenweitergabe noch wichtiger. So zeigen jüngste Entwicklungen wie das Beispiel betriebsinterner Fitnessprogramme in den USA, dass die Grenze zwischen angestoßenem und aufgezwungenem Tracking eine fließende ist. Insofern muss davon ausgegangen werden, dass in Zu-

kunft Personen zwar theoretisch weiterhin die Möglichkeit haben, auf die Verwendung von Wearables und auf das Self-Tracking ihrer Gesundheits- und Fitnessdaten zu verzichten.

Mit einem Verzicht könnten aber so große finanzielle oder anderweitige Nachteile verbunden sein, dass es praktisch unmöglich wird, sich der Nutzung dauerhaft zu entziehen. Bei einem Mitarbeiter einer Firma, der sich als Einziger weigert, an dem vom Arbeitgeber eingerichteten Programm zur Steigerung der Gesundheit der Belegschaft durch Tragen eines Fitnessarmbands teilzunehmen, erscheint dies aufgrund des erzeugten sozialen Drucks selbst dann evident, wenn die Firma keine direkten Sanktionen an die Nichtbefolgung knüpft.

Auch für die Akteure des Versicherungssektors werden analoge Entwicklungen erwartet.¹⁰² Die mit dem Einsatz von Big Data potenziell verbundene schleichende Entsolidarisierung wurde bereits angesprochen. Ganz konkret haben Versicherungen auch in Deutschland begonnen, Produkte anzubieten, welche die Tracking-Daten ihrer Kunden in die Preisgestaltung einbeziehen – wenn auch bislang nur in Form von in Aussicht gestellten Gutscheinen und Rabatten.¹⁰³

Für die Zukunft erscheint es aber keineswegs undenkbar, dass Verweigerer ganz konkret höhere Versicherungsbeiträge zahlen müssen als Kunden, die Wearables nutzen und ihre Daten bereitwillig an die Versicherung übermitteln. Noch extremere Szenarien werden für den Markt der Lebensversicherungen vorausgesagt. So äußerten sich Repräsentanten des Schweizer Rückversicherungsunternehmens Swiss Re kürzlich dahingehend, man müsse damit rechnen, dass es innerhalb der nächsten zehn Jahre unmöglich werden könnte, ohne das Tragen eines Wearables noch eine Lebensversicherung abzuschließen.¹⁰⁴

99 W. Nicholson Price II, „Black-Box Medicine“, *Harvard Journal of Law and Technology*, Vol. 28, 2015.

100 PWC, „Wearables“, s.o., S. 13.

101 Joseph Walker, „Can a Smartphone Tell if You're Depressed?“, *The Wall Street Journal Online*, 5. Januar 2015, online: <http://www.wsj.com/articles/can-a-smartphone-tell-if-youre-depressed-1420499238>.

102 Vgl. Michael Linden, „Generali will Fitnessdaten von Versicherten“, *golem.de*, 21. November 2014, online: <http://www.golem.de/news/telemonitoring-generalis-will-fitnessdaten-von-versicherten-1411-110722.html>.

103 Anne-Christin Gröger, „Generali erfindet den elektronischen Patienten“, *Süddeutsche Online*, 21. November 2014, online: <http://www.sueddeutsche.de/geld/neues-krankenversicherungsmodell-generalis-erfindet-den-elektronischen-patienten-1.2229667>.

104 Matthew Allen, „Versicherungen und die ‚Big-Data-Revolution‘“, *swissinfo.ch*, 22. April 2015, online: http://www.swissinfo.ch/ger/ohne-wearables-keine-krankenversicherung-_versicherungen-und-die-big-data-revolution-/41389092.

„Ich glaube nicht, dass Gesundheitsdaten per se private Daten sind. Sie sind Teil einer gesellschaftlichen Vorstellung und Transaktion und spiegeln wider, wie wir in der Gesellschaft mit Gesundheit umgehen. Dennoch dürfen diese Daten nicht zur illegitimen Diskriminierung führen.“

Lorena Jaume-Palasi, Direktorin für Kommunikation und Jugendbeteiligung am Europäischen Dialog für Internet Governance (EuroDIG) und Dozentin, 2. Expertenworkshop, 15.06.2015

„Es ist denkbar, dass Versicherungskosten für die Versicherung und die Versicherten selbst reduziert werden können, wenn diese ihren Health Score mit der Öffentlichkeit teilen.“

Peter Ohnemus, Präsident, Gründer und CEO dacadoo AG, Konsultation

dass die von Wearables erzeugten Daten nur so lange sicher sein werden, bis es sich für Kriminelle wirtschaftlich lohnt, sie zu hacken.

„Es gibt die gesellschaftliche Vereinbarung, dass bestimmte Daten nicht genutzt werden sollten. Der Arzt vermittelt der Krankenkasse beispielsweise nicht das genaue Krankheitsbild. Wir müssen diskutieren, ob wir in bestimmten Bereichen nicht auch ein ‚Nichtvorhandensein von Daten‘ brauchen.“

Stephan Noller, Gründer und Geschäftsführer der ubirch GmbH, 2. Expertenworkshop, 15.06.2015

Ein weiterer Aspekt des Schutzes der aufgezeichneten Daten ist im Hinblick auf eine mögliche staatliche Nutzung zu beachten. Dazu gehört nicht nur die auf der Hand liegende Problematik, dass die Datensätze von Geheimdiensten abgegriffen werden könnten, um beispielsweise Bewegungsprofile von Zielpersonen zu erstellen. Auch die Verwendung solcher Datensätze für die Justiz gewinnt an Bedeutung. So sind inzwischen einige Fälle bekannt, in denen die von Tracking-Apps generierten Daten in gerichtlichen Verfahren als Beweismittel zugelassen wurden. Schon 2013 wurden die durch einen Fitness-Tracker gespeicherten Informationen im Strafverfahren gegen einen Radfahrer in San Francisco ausgewertet, der einen tödlichen Unfall mit einem 71-jährigen Fußgänger verursacht hatte. Auf diese Weise wurde nachgewiesen, dass er vor der Kollision zu schnell gefahren war und mehrere rote Ampeln missachtet hatte.¹⁰⁷ Während es in jenem Verfahren noch um die Aufzeichnungen aus einer Smartphone-App ging, wurden im Herbst 2014 in Kanada zum ersten Mal von einem Wearable generierte Daten als Beweismittel zugelassen. Das Fitbit-Armband wurde von der Klägerseite ins Verfahren eingebracht, um die negativen Folgen eines Unfalls für die Gesundheit des Klägers aufzu-

Das Tracken von Gesundheitsdaten mittels Wearable ist stets mit einem Kontrollverlust verbunden. Selbst wenn ein Nutzer Daten nur für private Zwecke einsetzen möchte und deshalb die Weitergabe verweigern will, werden die aufgezeichneten Werte zumeist trotzdem in der Cloud, also auf einem Server, gespeichert. Ist dies aber erst einmal geschehen, gibt es keine technische Garantie mehr, dass die Daten nicht doch vom Anbieter des Geräts selbst oder einem Dritten ausgewertet werden. Dabei können sie im Zweifel auch zu kommerziellen Zwecken verwendet werden. Zudem sind die Daten auf dem externen Speicher anfällig für Kriminelle. Studien zufolge verschlüsseln die meisten Anbieter aktueller Wearables die übertragenen Daten gar nicht oder nur unzureichend.¹⁰⁵ Aber schon die Geräte selbst werden von Experten als nicht sicher eingestuft. Versuche haben gezeigt, dass es bei den meisten Geräten keineswegs schwierig ist, unbefugt an die aufgezeichneten Daten heranzukommen.¹⁰⁶ Aufgrund dieses Befundes wird erwartet,

105 Lupton 2016, s.o., S. 12.

106 Roman Unuchek, „How I hacked my smart bracelet“, Securelist, 26. März 2015, online: <https://securelist.com/blog/research/69369/how-i-hacked-my-smart-bracelet/>.

107 Karen Gullo, „San Francisco Cyclist Pleads Guilty to Manslaughter“, Bloomberg, 23. Juli 2013, online: <http://www.bloomberg.com/news/articles/2013-07-23/san-francisco-cyclist-pleads-guilty-to-manslaughter>.

zeigen. Dennoch wurde im Zusammenhang mit der Entscheidung des Gerichts von Kommentatoren angemerkt, es sei davon auszugehen, dass solche Beweiserhebungen bald die Norm auch in anders gelagerten Verfahren werden könnten.¹⁰⁸

Risiken des Einsatzes von Wearables in der medizinischen Forschung

Beim Einsatz von Wearables zum Zwecke der Durchführung medizinischer Studien entsteht ebenfalls eine Reihe problematischer Konstellationen. Da wäre zunächst erneut die Frage nach der informierten Zustimmung von Studienteilnehmern. Mehr noch als bei der Einwilligung in die Nutzung von Daten zu anderen Zwecken ist dies eine Grundvoraussetzung für die Zulässigkeit medizinischer Forschung mit menschlichen Probanden. Fehlt die Zustimmung ganz, oder wurde sie eingeholt, ohne dass die Person im Vorhinein umfassend über die Studie und die möglichen Folgen der Datenweitergabe informiert worden war, so bedeutet dies einen gravierenden Verstoß gegen medizinethische Prinzipien, da dann die Teilnehmer zum bloßen Mittel zum Zweck degradiert und somit in ihrer Würde betroffen sind. Probanden müssen stets als autonome Person behandelt werden, weshalb das Bemühen um eine informierte Einwilligung unerlässlich ist.¹⁰⁹ Gerade diese strenge Voraussetzung aber kann bei der Durchführung von Studien mittels Wearables und verknüpfter Apps nur schwerlich erfüllt werden. Selbst wenn Nutzer in die jeweilige Nutzung der Daten einwilligen müssen, wird fraglich bleiben, ob jeder Einzelne die Tragweite einer solchen Entscheidung tatsächlich hinreichend einschätzen kann. Bei der traditionellen medizinischen Forschung mit Probanden ist vorgeschrieben, dass jeder Teilnehmer mittels eines aufwendigen Prozesses über mögliche Risiken persönlich aufgeklärt wird.¹¹⁰ Das ist nicht gewährleistet, wenn die Zustimmung in der App am Bildschirm erfolgt, wie beispielsweise bei den Anwendungen auf Basis von Apples ResearchKit vorgesehen.

„Im Zusammenhang mit Big-Data-Geschäftsmodellen oder -Forschungen sollte die Transparenz für den einzelnen Nutzer möglichst umfassend sein.“

Dr. Alexander Dix, LL.M., Berliner Beauftragter für Datenschutz und Informationsfreiheit, 1. Expertenworkshop, 26.02.2015

Die bereits genannten Probleme bezüglich Datenschutz und Datensicherheit stellen sich auch dann, wenn die Wearables für medizinische Studien verwendet werden. So hat Apple zwar öffentlich erklärt, selbst weder Zugriff auf die auf Basis von ResearchKit-Anwendungen generierten Daten zu haben, noch sich einen solchen für die Zukunft vorzubehalten. Dennoch wurde die Frage aufgeworfen, ob nicht zumindest an kommerzieller Verwertung interessierte Parteien die Daten für ihre Zwecke würden nutzen können.¹¹¹ Zudem wurde wiederholt angemerkt, dass die zugesicherte Anonymisierung aller erzeugten personenbezogenen Daten möglicherweise nicht ausreichend sei, um zuverlässigen Datenschutz zu gewährleisten. Es bestehe stets die Gefahr, dass einzelne Studienteilnehmer anhand der übermittelten Werte doch identifiziert werden könnten.¹¹²

„In den Debatten zum Potenzial und zu den Risiken von Big Data werden dringend mehr konkrete Fallstudien benötigt, die eine differenzierte Sicht ermöglichen.“

Prof. Dr. Johannes Buchmann, Vizedirektor von CASED, Sprecher des Forschungszentrums CYSEC und des Sonderforschungsbereichs CROSSING und Professor für Informatik und Mathematik an der TU Darmstadt, 1. Expertenworkshop, 26.02.2015

108 Samuel Gibbs, „Cours sets legal precedent with evidence from Fitbit health tracker“, The Guardian Online, 18. November 2014, online: <http://www.theguardian.com/technology/2014/nov/18/court-accepts-data-fitbit-health-tracker>.

109 Mark A. Rothstein, „Ethical Issues in Big Data Health Research“, Journal of Law, Medicine and Ethics, Vol. 43, No. 2, 2015, online: <http://ssrn.com/abstract=2535373>.

110 Ebd.

111 Arielle Duhaime-Ross, „Apple’s new ResearchKit: ‘Ethics quagmire’ or medical research aid?“, The Verge, 10. März 2015, online: <http://www.theverge.com/2015/3/10/8177683/apple-research-kit-app-ethics-medical-research>.

112 Ebd.

Weiterhin ist nicht geklärt, wie fehleranfällig Erhebungen mittels Wearables in der Praxis sind. Nutzer können bei Anwendung der Apps nicht wie bei traditionellen medizinischen Studien überwacht werden, weshalb von einer erhöhten statistischen Ungenauigkeit auszugehen ist. Das gilt selbst dann, wenn sicher ist, dass die Sensoren der Geräte selbst hinreichend zuverlässig messen.¹¹³ Das größere Problem werden aber voraussichtlich dennoch die Nutzer selbst darstellen. Wenn es keinen persönlichen Kontakt zu den Probanden gibt, kann nicht ausgeschlossen werden, dass Teilnehmerdaten von vornherein verfälscht sind. Es gibt keinen Prüfmechanismus, um festzustellen, ob Grunddaten wie beispielsweise das Alter der Person wahrheitsgemäß eingegeben wurden. Aus diesem Umstand ergeben sich zudem rechtliche Probleme. So ist es gesetzlich verboten, Minderjährige ohne Einverständniserklärung ihrer Eltern an medizinischen Studien teilnehmen zu lassen. Auch hier fehlt aufgrund des Versuchsaufbaus die Möglichkeit, das Einhalten der Altersgrenze zu verifizieren.

Das Forschungsergebnis kann schließlich durch die sogenannte Stichprobenverzerrung gestört oder im Extremfall sogar gänzlich unbrauchbar sein. Werden medizinische Daten für eine bestimmte Studie ausschließlich von Personen gesammelt, die im Besitz eines Wearables sind, dann schließt das all jene aus, die sich ein solches (ggf. teures) Gerät nicht leisten können. Schlussfolgerungen aus den präsentierten Studien können daher dazu führen, dass die Interessen, Bedürfnisse und spezifischen Eigenschaften der nicht in die Datenbasis eingeflossenen Teile der Bevölkerung nur unzureichend berücksichtigt werden. Diesen Bedenken wird allerdings zugleich entgegengehalten, mit traditionellen Methoden durchgeführte Studien seien ebenfalls nicht vor solch einer Verzerrung gefeit. Auch sie hätten zumeist nur Zugriff auf einen bestimmten Teil der Bevölkerung. Ihre Probanden lebten überproportional häufig in der Nähe von Universitäten, mithin in urbanen Zentren,

und hätten genügend Zeit für die Teilnahme. Daher könne man sogar von einer Erhöhung der Repräsentativität ausgehen, sobald sich Wearables erst einmal so weit am Markt etabliert haben, dass sie so selbstverständlich sind wie heute schon Smartphones.

Schlussendlich kann sich der vermehrte Einsatz von Wearables auf das Verhältnis zwischen Arzt und Patient auswirken. Insofern die durch das Gerät generierten Daten genutzt werden, um Diagnosen zu erstellen und Therapiepläne zu entwickeln, führt die Nutzung der Technologie zu einer zunehmenden Entpersonalisierung der medizinischen Direktversorgung –, die Kommunikation mit dem Patienten könnte leiden, wenn Algorithmen darüber entscheiden, wie in Bezug auf den Gesundheitszustand weiter verfahren wird. Das kann unter Umständen so weit gehen, dass der Arzt die algorithmusbasierten Diagnosen selbst kaum noch verstehen oder nachvollziehen kann und deshalb auch nicht mehr in der Lage ist, den Patienten seiner Pflicht entsprechend umfassend zu informieren.

Die Herausforderungen für die Gesellschaft

Das Tracking bei Smart-Health-Anwendungen wirft noch wesentlich gravierendere gesellschaftliche Fragen auf als Tracking im Fall von Smart Mobility. Gesundheitsdaten sind beinahe per definitionem höchstpersönlicher Natur. Hersteller und Anbieter von Geräten und Anwendungen, deren Zweck es ist, aus solchen Werten Big Data zu generieren und zu verwerten, müssen diese sensible Qualität unbedingt und stets berücksichtigen. Dabei ist es zunächst einmal zweitrangig, ob die Daten für kommerzielle oder wissenschaftliche Zwecke genutzt werden sollen. Die im Folgenden angerissenen Fragen in Bezug auf Smart Health sind als Grundlage für die Erarbeitung und Aushandlung einzelner Regelungselemente gedacht, die in einem weiteren Schritt in einen möglicherweise aufzustellenden Digitalen Kodex für den Umgang mit Big Data im Gesundheitsbereich überführt werden könnten.

113 Stephanie M. Lee, „Why Apple’s New ResearchKit Could Have a Diversity Problem“, BuzzFeed News, 10. März 2015, online: <http://www.buzzfeed.com/stephaniemlee/why-apples-new-researchkit-could-have-a-diversity-problem#.npN60yG19>.

„Wie beantwortet eine Gesellschaft, in der eine unglaubliche Menge an Daten zirkuliert, die Frage, ob es autonome informationelle Selbstbestimmung geben kann, geben soll oder geben muss?“

Christian Hawellek, wissenschaftlicher Mitarbeiter am Institut für Rechtsinformatik der Leibniz Universität Hannover, 1. Expertenworkshop, 26.02.2015

Der Aspekt der Freiwilligkeit ist von entscheidender Bedeutung: einerseits bezüglich der Nutzung von Wearables, andererseits bezüglich der Weitergabe der mittels solcher Geräte erzeugten Daten. Hier ist die Selbstbestimmung der Nutzer betroffen. Es muss erörtert werden, ob und inwieweit die Entscheidung, solche Geräte zu nutzen, zur individuellen Verfügung gestellt bleiben soll. Wenn es dabei bleiben soll, dass niemand gezwungen werden kann, seine Gesundheitsdaten Versicherungen, Arbeitgebern oder Unternehmen preiszugeben, dann muss diskutiert werden, wie dieses Prinzip auch gegen steigenden Druck von Seiten dieser interessierten Akteure durchgesetzt werden kann.

Dazu gehört die Frage, was mit Personen geschieht, die sich der Technologie verweigern. Inwieweit müssen diese daraus folgende Nachteile zum Beispiel bei ihrem Versicherungsmodell hinnehmen, und von welchem Punkt an müssen sie vor unverhältnismäßiger Benachteiligung durch Versicherungen oder andere Akteure geschützt werden?

Darüber hinaus stellt sich die Herausforderung, jene Personen vor Diskriminierung zu schützen, die ihre Daten sehr wohl teilen, aufgrund algorithmischer Kalkulationen von Wahrscheinlichkeiten aber einer marginalisierten gesellschaftlichen Gruppe zugeordnet werden, die systematischer Benachteiligung ausgesetzt ist. Gerade hinsichtlich stigmatisierter Krankheitsbilder muss diese Möglichkeit frühzeitig erörtert und kritisch hinterfragt werden. Bei Versicherungsmodellen, die dank präziser Datenerhebung durch Tracking an Eigenschaften und Verhaltensweisen des Versicherungsnehmers anknüpfen können, ist eine Preisdifferenzierung inhärent und gewollt. Das gilt für Krankenversicherungen ebenso wie für Kfz-Versicherungen. Die Frage muss aber stets sein, wann gerechtfertigte und es-

senzielle Differenzierung in nicht mehr zu rechtfertigende Diskriminierung umschlägt.

„Die Individualisierung von Krankenkassenbeiträgen ist mit dem Solidaritätsprinzip nicht zu vereinbaren und daher auch für uns als gesetzliche Krankenkassen kein Thema.“

Dr. Kai Kolpatzik, Leiter Prävention beim AOK-Bundesverband, 2. Expertenworkshop, 15.06.2015

Andererseits muss diskutiert werden, ob die Nutzung von Big Data im Krankenversicherungssektor nicht geradezu notwendig ist, um die Finanzierbarkeit des Gesundheitssystems zu sichern. Wenn das System nur durch Preisdifferenzierung tragfähig gehalten werden kann, ließe sich argumentieren, dass die allgemeine Offenlegung aller Gesundheitsdaten Voraussetzung ist, um die Solidargemeinschaft aufrechtzuerhalten, und deshalb verpflichtend erfolgen müsse. Die Frage nach der Freiwilligkeit verlöre unter dieser Prämisse an Gewicht. Diese könnte aber zu einer systematischen Benachteiligung von Randgruppen führen.

„Krankenversicherungen sind eine wichtige Grundversorgung. Sie sollten für alle gleich viel kosten, unabhängig von den persönlichen Risikofaktoren.“

Julia Reda, MdEP, Mitglied im Rechtsausschuss des EP und Berichterstatterin für die Evaluation der Umsetzung der Urheberrechtsrichtlinie von 2001, Konsultation

Ungerechtfertigte Diskriminierung kann im Zweifel nur entdeckt werden, wenn die Algorithmen, die den Bewertungen einzelner Nutzer zugrunde liegen, offengelegt werden. Die sich anschließende Frage ist daher die nach Transparenzpflichten für Versicherungen und andere im Bereich Big Data tätige Akteure. Das Recht, auf Antrag Berechnungsgrundlagen einsehen zu können – soweit sie realistischerweise verständlich gemacht werden können –, könnte dazu beitragen, Manipulationen oder sonstige negative Effekte zu verhindern.

Ein solcher Anspruch könnte jedoch nur den ersten Schritt darstellen. Als Nächstes wäre zu erörtern, ob Kunden oder andere betroffene Personen, die mit ihren persönlichen Daten das Rohmaterial für algorithmische Festlegungen bereitstellen, die Berechnungsmodelle anfechten könnten, wenn diskriminierende oder andere Tendenzen offenbar werden, die sie ungerechtfertigt benachteiligen. Eine andere, eventuell realistischere und effizientere Möglichkeit, um diese Problematik anzugehen, wäre die Einrichtung einer neutralen Instanz, die für die Prüfung von Algorithmen zuständig wäre – eine Art „Algorithmen-TÜV“ – und die im Zweifelsfall die Verwendung diskriminierender oder anderer unangemessen benachteiligender Verfahren untersagen könnte.

Unmittelbar mit diesen Fragen ist das Problem der informierten Zustimmung verbunden. Freiwilligkeit kann nur dann gewährleistet sein, wenn Nutzer in ausreichendem Maße über Voraussetzungen, Funktionsweise und vor allem Konsequenzen des Trackings und der Übermittlung von Gesundheitsdaten informiert sind. Nur von einer soliden Wissensbasis aus kann eine wirklich autonome Entscheidung auf diesem komplexen Gebiet getroffen werden. Auch hier könnte eine neutrale Überprüfungsinstanz Abhilfe schaffen.

„Durch die unbedachte Übermittlung von Gesundheitsdaten, die nach unbekanntem Algorithmen ausgewertet werden, entsteht für den Versicherten ein unkalkulierbares datenschutzrechtliches Risiko.“

Andrea Voßhoff,
Bundesbeauftragte für den Datenschutz
und die Informationsfreiheit,
Interview

Ein davon unabhängiger Problemkomplex, der dennoch gesellschaftlicher Klärung bedarf, ist die Infragestellung traditioneller Grundprinzipien medizinischer Forschung in Anbetracht der Möglichkeit, Studien mittels Wearables durchzuführen. Hier ist zwar ebenfalls die Frage nach der informierten Zustimmung der Probanden berührt. Darüber hinaus besteht die Gefahr der Stichprobenverzerrung und das Problem der Genauigkeit und Zuverlässigkeit der ermittelten Daten – Aspekte, die bei der weiteren Entwicklung bedacht werden müssen. Wenn die Chancen der umfassenden, ständigen Datenerhebung mittels neuer Technologien von der Forschung adäquat genutzt werden sollen, muss erörtert werden, ob und inwieweit die hergebrachten Prinzipien medizinischer Forschung einer Anpassung oder Neuausrichtung bedürfen.

4. Mögliche Kodifizierungen im Bereich Big Data

4.1 Übergreifende Aspekte der Anwendungsfelder

Die Beschreibung von Smart Mobility und Smart Health anhand der drei Beispiele telematikbasierte Autoversicherung, Verkehrslenkung und Wearables in den vorangegangenen Kapiteln hat viele spezifische Aspekte, aber auch Gemeinsamkeiten zutage gefördert. Letztere liegen zumeist in der Technologie des Trackings zur Datenerhebung selbst begründet. Die Methode erfasst stets vom Nutzer erzeugte Daten, die persönlicher als auch allgemeiner Natur sind. Während das Tracking von Ortsdaten für die Optimierung der Verkehrsströme in Großstädten vergleichsweise wenig in die Autonomie des Individuums eingreift, sind die Gesundheitsdaten, die von Fitnessarmbändern und Smart Watches erfasst werden, stets höchstpersönlicher Natur. Allen Anwendungen ist gemeinsam, dass die gesammelten Daten nach ihrer Zusammenführung potenziell für Zwecke verfügbar sind, die über das hinausgehen, was ursprünglich mit den Nutzern vereinbart oder ihnen angekündigt wurde. Jede dieser Datenerhebungen ist insoweit mit einem Eingriff in die Privatsphäre verbunden, dass die Auswertung der Daten Rückschlüsse über das Leben des Einzelnen zulässt – seien es seine Ernährungsweise oder Bewegungsprofile. Big Data macht aufgrund der Einbeziehung großer Datenmengen im Zusammenspiel mit auf Wahrscheinlichkeitsrechnungen basierenden Algorithmen darüber hinaus Prognosen über Individuen sogar dann möglich, wenn deren Da-

ten für sich genommen pseudonymisiert oder anonymisiert erfasst worden waren. Big-Data-Analysen differenzieren und diskriminieren. Das ist ihre Essenz.

Ein anhand von gesellschaftlichen Aushandlungsprozessen erarbeiteter Digitaler Kodex für den Umgang mit Big Data muss diese Aspekte berücksichtigen. Zugleich darf er aber die Chancen und Potenziale, die mit der Nutzung von Big Data verbunden sind, nicht übermäßig beschränken oder gar ersticken. Ein Digitaler Kodex sollte zum Ziel haben, dass die Chancen der neuen Technologien zum allgemeinen gesellschaftlichen Vorteil genutzt werden können. Er sollte deshalb Innovationen in diesem Bereich gerade nicht ausbremsen oder erschweren, sondern fördern. Ein zusammen mit den relevanten Akteuren diskutiertes und ausgehandeltes Regelwerk, das am Markt als handlungsleitend akzeptiert und befolgt würde, könnte insbesondere die Chancen kleinerer Anbieter von Datenanwendungen, die auf Tracking basieren, erhöhen, sich gegenüber Mitbewerbern durchzusetzen. Ein rücksichtsvoller Umgang mit den Daten der Nutzer, der durch die Existenz eines Kodex gewährleistet und abgesichert wird, ist von dieser Perspektive aus betrachtet weniger ein Hemmnis für Innovationen als vielmehr ein förderndes Moment, indem bei den Nutzern Vertrauen geschaffen wird und Unsicherheiten bei der Verwendung der bereitgestellten persönlichen Daten verringert werden. Ein Kodex würde so zu einem positiven Anreizsystem für Big-Data-Anwendungen. Dabei ist es zunächst einmal zweitrangig, ob es am Ende einen umfassenden Digitalen Kodex für Big

Data gäbe, der alle in diesem Bericht genannten Aspekte abdeckt, oder ob für Unterthemen, z.B. Smart Mobility oder Smart Health, eigene Kodizes formuliert würden. Welcher Weg hier am sinnvollsten ist, müssten Diskurs und Aushandlungsprozess zeigen.

„Viel stärker als juristische Regelungen wirken oft soziale Normen – und immer häufiger werden diese von einer globalen Denkkultur geprägt. Staatliche Reaktionen erscheinen da oft unpassend, ein ausgehandelter Prozess mit allen Stakeholdern im internationalen Rahmen erscheint hier sinnvoller.“

Tobias Schwarz, Leitender
Chefredakteur Netzpiloten.de,
1. Expertenworkshop, 26.02.2015

4.2 Grundkonflikte

Im Verlauf des Berichts sind neben den Potenzialen von Big-Data-Anwendungen im Bereich Smart Mobility und Smart Health mehrere Konfliktlinien sichtbar geworden, die an verschiedenen Stellen bereits angedeutet wurden. Diese Grundkonflikte beziehen sich auf problematische Aspekte der Erhebung und Auswertung von Datenbeständen, die bereits das geltende traditionelle Datenschutzrecht adressiert. Das gilt insbesondere für das Problem der Profilbildung, welches aus der umfassenden Digitalisierung resultiert. Daten sammelnde Unternehmen können ihre Datenbestände, zum Beispiel über das Surfverhalten von Nutzern, auf einzelne Personen zurückführen. Über Korrelationen können sie dann Wahrscheinlichkeitsaussagen zu Fragen treffen, die den Nutzer in sensiblen Bereichen betreffen, ohne dass er dazu jemals direkt Auskunft gegeben hat. Es ist sogar denkbar, dass aus den Daten auf Informationen geschlossen wird, die über das hinausgehen, was das Individuum über sich selbst weiß. Es können beispielsweise Aussagen über das Vorliegen von unerkannten Krankheiten oder von sexuellen Vorlieben auf Grundlage des Verhaltens in Sozialen Netzwerken extrapoliert werden. Hinzu kommt, dass in solchen Fällen mitunter nicht nur über die Mitglieder des Sozialen Netzwerks selbst

Aussagen möglich sind, sondern auch über Dritte, die nicht bei dem Netzwerk angemeldet sind und es vielleicht nie waren.

„Wird Big Data als Grundlage für Steuerungsmaßnahmen verwendet, eröffnet dies immer ethische Fragen – unter anderem nach Privatheit und Sicherheit.“

Lorena Jaume-Palasi, Direktorin
für Kommunikation und
Jugendbeteiligung am Europäischen
Dialog für Internet Governance
(EuroDIG) und Dozentin,
1. Expertenworkshop, 26.02.2015

Dass hier die Privatsphäre und das Persönlichkeitsrecht der entsprechenden Personen betroffen sind, ist offensichtlich. Bereits ausgeführt wurde, dass das Datenschutzrecht im Hinblick auf Big Data an die Grenzen seiner Steuerungsfähigkeit gelangt. Da es von bewussten Entscheidungen der handelnden Akteure ausgeht, muss es dort scheitern, wo entweder keine Einwilligung zur Datenverarbeitung mangels Wissen um die Erhebung vorliegt oder wo die Folgen der Einwilligung für den Einzelnen nicht absehbar sind. Dort, wo das Datenschutzrecht nicht greift, kann ein ausgehandelter Digitaler Kodex ansetzen. Er kann zur Klärung der Frage beitragen, wie verhindert werden kann, dass die gesammelten Daten zweckentfremdet beziehungsweise über die vereinbarten Zwecke hinaus genutzt werden. Zudem könnte er regulieren, welche Anforderungen an Anonymisierung und Pseudonymisierung zu stellen sind, wenn die Daten nicht nur für die Erfüllung gegenseitiger Vertragsverhältnisse verwendet werden sollen.

Neben der Frage nach dem Prinzip der Zweckbindung der Datenerhebung, das durch die explorative Ausrichtung von Big Data unterminiert wird, könnte mit dem Kodex auch der Aspekt der Datensparsamkeit abgedeckt werden. Big-Data-Anwendungen sind ihrem Wesen nach auf große Datenmengen angewiesen. Trotzdem bleibt die Frage sinnvoll und berechtigt, ob der „Sammelwut“ nicht Grenzen gesetzt werden sollten. Gerade wenn die Erfassung, Speicherung und Übermittlung persönlicher Daten im Rahmen der Nutzung von Apps oder anderer Software völlig

zweckfrei erfolgt, ist die Datenerhebung unter Umständen als nicht mehr verhältnismäßig zu qualifizieren. Nutzerdaten sind in den vergangenen Jahren zu einem immer bedeutenderen Wirtschaftsgut geworden. Es hat sich ein Sekundärmarkt für den Handel mit Daten entwickelt, der von den primären Anwendungen, durch die die Daten überhaupt erst generiert werden, weitgehend unabhängig geworden ist. Das ist risikobehaftet, denn dieser Umstand erzeugt bei den Unternehmen immer größere Anreize, so viele Daten wie möglich zu sammeln. Bekanntheit erlangte in diesem Zusammenhang das Beispiel einer Taschenlampen-App für das iPhone. Sie war kostenfrei erhältlich, mit der Installation aber willigten die arglosen Nutzer in die Übermittlung einer Vielzahl sehr persönlicher Daten an den App-Anbieter ein. Ein Kodex könnte in solchen Fällen dazu beitragen, diese Art von Geschäftspraxis entweder ganz zu unterbinden oder jedenfalls transparenter zu machen.

Daneben zeigt sich eine potenzielle Konfliktlinie zwischen jenen Nutzern, die der Verwendung ihrer persönlichen Daten aus freien Stücken zustimmen,

weil sie sich davon bestimmte Vorteile versprechen, und jenen, die sich aus Datenschutz- oder sonstigen Erwägungen dagegen entscheiden. Zum einen entsteht gesellschaftlicher Druck auf diejenigen, die eigentlich nicht mitmachen wollen. Zum anderen können die Daten, die von Personen freiwillig überlassen wurden, diejenigen negativ beeinflussen, die sich der Herausgabe ihrer Daten verweigern. Hier könnte ein Kodex Grundsätze formulieren, wie beispielsweise unverhältnismäßige Benachteiligungen für die Gruppe der Verweigerer zu verhindern wären.

„Wir brauchen grundsätzlich neue Regeln. Dabei sollten allerdings die alten Werte wie etwa das Solidarprinzip oder die Datensparsamkeit nicht über Bord geworfen werden, sondern in smarte neue Leitlinien überführt werden.“

Stephan Noller, Gründer und Geschäftsführer der ubirch GmbH, 2. Expertenworkshop, 15.06.2015

INTERVIEW MIT TOBIAS O. KEBER

„Privacy made in Germany“ könnte ein Verkaufsargument sein

? Gibt es aus Ihrer Sicht einen qualitativen Unterschied zwischen Datensammlungen im analogen Zeitalter, die es seit Jahrzehnten gibt, und den neuen Datensammlungen im digitalen Zeitalter? Also zum Beispiel dem Kredit scoring durch die

Schufa und dem Tracking und der Profilbildung im Internet.

Tobias Keber: Um den Unterschied zu erklären, können wir uns die Frage stellen: Was erwartet der Nutzer? Versteht er, was passiert? Bei der Schufa ist es nachvollziehbar, dass irgendwo gespeichert

wird, ob jemand einen Kredit aufgenommen hat, wie hoch er ist und ob er ihn zurückzahlt. Man versteht, dass die gegenwärtigen und zukünftigen Gläubiger ein Interesse daran haben, dass diese Informationen verfügbar sind. Die Schufa muss sich aber auch an die Regeln halten, die im Bundesdatenschutzgesetz stehen.

Prof. Dr. Tobias O. Keber

Prof. Dr. iur. Tobias O. Keber, seit Oktober 2012 Professor für Medienrecht und Medienpolitik in der digitalen Gesellschaft, Hochschule der Medien (HdM) Stuttgart, daneben Lehrbeauftragter für Internet- und Multimediarecht am Mainzer Medieninstitut (Johannes Gutenberg-Universität Mainz) sowie an der Universität Koblenz-Landau. Keber ist Vorsitzender des Wissenschaftlichen Beirats der Gesellschaft für Datenschutz und

Datensicherheit (GDD), Herausgeberbeirat der Zeitschrift „Recht der Datenverarbeitung“ (RDV), Zeitschrift für Datenschutz-, Informations- und Kommunikationsrecht, sowie im Leitungsgremium des Instituts für Digitale Ethik (IDE). Prof. Dr. iur. Tobias Keber war vor seiner akademischen Laufbahn als Rechtsanwalt tätig. Er ist Autor zahlreicher Fachpublikationen zum nationalen und internationalen Medien-, IT- und Datenschutzrecht.

Foto: A. Krombholz



„PRIVACY MADE IN GERMANY“ KÖNNTE EIN VERKAUFSARGUMENT SEIN

Beim Web-Tracking und anderen digitalen Formen des Datensammelns ist grundsätzlich das Problem, dass man sich nicht immer bewusst ist, dass und wie genau hier Profile erstellt werden. Der smarte Fernseher ist dafür ein gutes Beispiel. Ich erwarte nicht, dass es einen Rückkanal gibt, über den mein Nutzungsverhalten verfolgt wird. Beim klassischen Fernseher weiß ich, der Mensch, der da die Nachrichten spricht, sieht mich nicht. Das ist heute anders: Wir sitzen vor einem Fernsehcomputer. Unser Nutzungsverhalten kann aufgezeichnet und verarbeitet werden. Das haben die meisten Leute noch nicht auf dem Schirm. Das Bewusstsein dafür entsteht gerade erst.

Der andere interessante Punkt bei Big Data ist, dass Sie in der Lage sind, aus Daten etwas zu machen, von dem Sie zum Zeitpunkt der Datenerhebung noch gar nicht wussten, dass es möglich sein wird. Das ist ein Problem für unser

gegenwärtiges Datenschutzrecht. Das geht nämlich davon aus, dass ich nur Daten zu einem bestimmten Zweck erheben darf, über den im Grundsatz vorher informiert werden muss. Big Data ermöglicht es aber, neue Verknüpfungen zu finden, über die man vielleicht noch gar nicht nachgedacht hat. Das große Problem ist also, dass nicht vorhersehbar ist, welche Informationen aus den Daten zukünftig generiert werden können.

? **Wie kann man das verhindern? Brauchen wir mehr Transparenz- und Informationspflichten für Daten?**

TK: Der Transparenzgrundsatz ist ein wichtiger Grundsatz im Datenschutzrecht, und er hat auch nicht ausgedient. Wir brauchen einfache Mechanismen, die dem Nutzer zeigen: Da passiert etwas. Das muss nicht in jedem Fall bis ins letzte Detail gehen. Ich weiß

bei meinem Auto auch nicht, wie genau der Motor funktioniert, aber die Grundsätze müssen schon klar sein. Niemand liest überlange Datenschutzerklärungen bspw. bei Google und Facebook. Sie sind hochkomplex und unverständlich. Wir brauchen eine klare, einfache Sprache, und es gibt Vorschläge, das über Icons und Piktogramme zu lösen. Grundsätzlich muss es so gestaltet sein, dass es der durchschnittliche Nutzer verstehen kann. Dafür kann man sich ein mehrstufiges Verfahren vorstellen: Diejenigen, die nicht so viele Informationen brauchen, bekommen nur die Grundaussagen und geben auf dieser Grundlage ihre Zustimmung. Der Nächste will es vielleicht ein bisschen genauer wissen, und der Dritte will bei jeder Datenverarbeitung en détail informiert werden. Der Nutzer sollte auch insoweit die Wahl haben. Daneben und unabhängig davon beruht ein zukunftsfähiges, smartes Datenschutzrecht auf den

Der Transparenzgrundsatz ist ein wichtiger Grundsatz im Datenschutzrecht, und er hat auch nicht ausgedient.

Tobias O. Keber

Konzepten von Privacy by Design und Privacy by Default, d.h. dass der Schutz der Privatsphäre schon bei der Gestaltung und in den Voreinstellungen der Angebote eine Rolle spielt.

? Wie sehen Sie die Bereitschaft der Unternehmen, dabei mitzuziehen?

TK: Das ist ein Problem. Im Kreise der IKT-Interessenverbände ist bisweilen zu hören: „Datenschutz ist innovationshemmend, und die Amerikaner hängen uns ab.“ Dieser Einwand ist zum Teil berechtigt. Aber vielleicht muss man versuchen, den Spieß umzudrehen. Ein Siegel „Privacy made in Germany“ oder „Privacy made in Europe“ sollte Verkaufsargument werden. Geräte, die diese Auszeichnung tragen, könnten gekauft werden, gerade weil der Privacy-Standard so hoch ist. Man darf den Datenschutz nicht immer nur als Hemmschuh begreifen.

? In den letzten Jahren spricht man vermehrt vom Internet of Things. Immer mehr Geräte sind digital vernetzt und funken Daten zurück an die Hersteller und Softwareentwickler. Diese Geräte haben keine Benutzeroberfläche mehr, sodass es gar nicht möglich ist, sich mit einer Datenerhebung einverstanden zu erklären. Wie könnte man so etwas regeln?

TK: Da kann man über eine Kennzeichnungspflicht nachdenken – zum Beispiel Aufkleber, die deutlich machen: „Dieses Gerät verschickt Daten.“ Das Wissen, dass Geräte potenziell Daten versenden und verarbeiten, muss sich allerdings erst in der Gesellschaft durchsetzen. Technischer Fortschritt und neue Datenverarbeitungsmethoden sind nicht automatisch das Ende von Privatheit

und Datenschutz. Technik an sich ist weder gut noch böse, wenn Sie in ethischen Kategorien denken wollen. Es kommt auf ihren konkreten Einsatz an. Big Data hat zum Teil sehr sinnvolle Anwendungsmöglichkeiten. Sie können etwa bei einem Ebola-Ausbruch mit Big-Data-Analysen ziemlich genau sagen, wie die Epidemie verlaufen wird, und einschätzen, wo genau die Hilfe vornehmlich gebraucht wird.

? Sie haben gerade schon über Vereinfachungen bei den Nutzungsbedingungen und Einverständniserklärungen gesprochen. Gibt es noch andere Maßnahmen, die man treffen kann?

TK: Wir brauchen auf jeden Fall eine neue gesellschaftliche Diskussion über Datenschutz und Privatsphäre. Sie darf nicht erst bei Erwachsenen ansetzen, son-



„PRIVACY MADE IN GERMANY“ KÖNNTE EIN VERKAUFSARGUMENT SEIN



dem muss schon in der Schule und sogar im Kindergarten stattfinden. Dabei sollte es nicht nur um technische Kompetenzen gehen, sondern auch um ethische. Wir müssen uns dem Privatheits-Paradoxon stellen. Damit meine ich den Umstand, dass wir zwar erkennen, dass Privatsphäre ein wichtiges Gut ist, es im Alltag aber nicht umsetzen. Selbst nach den Snowden-Enthüllungen änderten die wenigsten User ihr Verhalten, dass sie zum Beispiel ihre Festplatten und ihre E-Mails verschlüsseln oder bestimmte unsichere Dienste nicht mehr benutzen. Trotzdem sagen die meisten Leute in den Umfragen, dass sie nicht überwacht werden wollen.

Die User wissen mittlerweile, dass Facebook nicht wirklich kostenlos ist, sondern dass sie mit ihren Daten bezahlen – wie auch bei anderen scheinbar kostenlosen Diensten im Internet. Aber wie wertvoll ihr Datensatz genau ist und was sie dafür be-

kommen könnten, das wissen sie nicht. Vielleicht ist das nämlich viel mehr als die schlichte Nutzungsmöglichkeit eines doch vergleichbar faden Sozialen Netzwerks. Kurz: Wir brauchen mehr Privatheitskompetenz.

? Sprechen wir über konkrete Beispiele: Das vernetzte Auto ist schon Realität. Verschiedene Akteure haben ein Interesse an den Daten, die so ein Auto generiert: Hersteller, Versicherungen, Softwareentwickler, womöglich noch Drittanbieter, die die Daten kaufen. Wo sehen Sie die ethischen Grenzen im Umgang mit den Daten?

TK: Problematisch wird es insbesondere bei Dritten, die nicht mehr unmittelbar mit dem Fahrzeug und seiner Herstellung in

Verbindung stehen. Dazu gehören etwa die Versicherer. Eine Datenerhebung durch sie geht nur mit ausdrücklicher Zustimmung. Aber selbst wenn diese vorliegt, ist man manchmal verwundert, für wie wenig Geld Leute bereit wären, ihr Fahrverhalten aufzeichnen und auswerten zu lassen. Wenn ich beispielsweise meine Zustimmung zum Tracking geben würde, nur um einen Bonus von 50 Euro im Jahr zu bekommen, dann wäre das eigentlich irrational. 50 Euro für einen Datensatz, in dem mein ganzes Fahrverhalten über ein Jahr gespeichert ist. Ernsthaft? Im Datensatz findet sich jeder Bremsvorgang, jedes Mal, wenn ich ordentlich Gas gegeben habe und wo ich mich und wann und wie lange aufgehalten habe. Dieser Datensatz dürfte ganz sicher wertvoller sein als der Bonus in meinem Beispiel.

? Wie wertvoll ist er denn genau?

“ **Man musste den Leuten erklären, was eine Währung ist und wie viel wert sie ist. Wenn Daten Währungen sind, muss man sich dazu mehr Gedanken machen.** Tobias O. Keber

TK: Ihre Frage trifft des Pudels Kern. Das weiß im Augenblick keiner. Es gibt verschiedene Ansätze, den (ökonomischen) Wert der Daten genauer zu bestimmen. Wenn Sie sich ein Facebook-Profil anschauen: Ist es die Anzahl der „Gefällt mir“-Angaben, was geliked wurde, wie viele Freunde die Likes weiterteilen? Hier steckt man zum Teil noch in den Kinderschuhen. Es gibt noch kein standardisiertes Verfahren. Das erinnert ein bisschen an die Situation, als man die Neue Welt erobert hat und mit den Indianern Glasperlen gegen Gold getauscht hat. Man musste den Leuten erklären, was eine Währung ist und wie viel wert sie ist. Wenn Daten Währungen sind, muss man sich dazu mehr Gedanken machen. Aber das findet kaum statt.

? **Im Gesundheitsbereich haben wir inzwischen die ersten Angebote von Versicherungen, dass Leute**

einen billigeren Versicherungstarif erhalten, wenn sie ihre Daten abgeben, die über Fitnessarmbänder oder Ähnliches gesammelt wurden. Wie bewerten Sie das?

TK: Ich sehe das sehr kritisch. Hier stellt sich die Frage, ob den Leuten überhaupt bewusst ist, wie sensibel Gesundheitsdaten, wie Blutdruck, Herzfrequenz, Blutzuckerspiegel etc. eigentlich sind. Wäre ein Bonus von 50 Euro es wirklich wert, dass diese Daten dazu genutzt werden, Prognosen über die persönliche Gesundheitsentwicklung anzustellen? Ein konkretes Beispiel: Ich habe einen solchen Versicherungstarif mit Gesundheitsdatenüberwachung, und plötzlich kündigt mir die private Versicherung, weil sich irgendetwas an meinen Gesundheitsdaten geändert hat. Diese Veränderung bedeutet statistisch, dass ich wahrscheinlich inner-

halb der nächsten fünf Jahre eine Herzerkrankung bekomme. Deswegen kündigt die Versicherung jetzt schon mal. Das heißt, durch diese Vorhersagen aufgrund von Big Data entstehen Folgen, die ich mit meiner Einwilligung so nicht voraussehen konnte. Unser Rechtssystem adressiert diese Konstellationen gegenwärtig nur unzureichend.

? **In Deutschland gibt es seit Jahrzehnten ein solidarisches Krankenversicherungssystem. Wenn man jetzt damit anfängt, Krankenversicherungen aufgrund der individuellen Daten zu vergeben, ist das nicht das Ende des jetzigen Systems?**

TK: Ja, das könnte zumindest der Anfang vom Ende der Solidargemeinschaft sein. Das Risiko würde nicht mehr auf die Schultern von vielen verteilt, sondern für jeden



„PRIVACY MADE IN GERMANY“ KÖNNTE EIN VERKAUFSARGUMENT SEIN

Die Frage ist, wie genau unsere Verfassung die Prinzipien Solidarität auf der einen Seite und Selbstverantwortlichkeit auf der anderen Seite austariert.

Tobias O. Keber




individuell berechnet. Das würde aber wohl auch bedeuten, dass Sie signifikante Einschränkungen ihrer persönlichen Freiheit hinnehmen müssten. Ohne unmittelbare finanzielle Folgen für Ihren Versicherungstarif könnten Sie nicht mehr über den Durst trinken, fett essen und auch keinem Risikosport mehr nachgehen. Wenn Sie also Fallschirmspringen gehen wollen, sagt Ihnen die Krankenversicherung: „Kein Problem, es wird aber teurer.“ Das wäre ein komplett neues gesellschaftliches Konzept, und das muss man nicht gut finden.

? Ist das ein Punkt, den eine Gesellschaft aushandeln und ausdiskutieren muss? Oder sehen Sie es als Pflicht des Gesetzgebers an, solche Prinzipien von Solidargemeinschaft zu verteidigen?

TK: Das ist eine verfassungsrechtlich sehr interessante Frage. Es gibt Schutzpflichten, denen der Gesetzgeber durch den Erlass von Normen Rechnung tragen muss. Das muss die Legislative gegebenenfalls auch gegen den Willen von Partikularinteressen tun. Andererseits muss der parlamentarische Gesetzgeber auch auf Veränderungen im gesellschaftlichen Diskurs reagieren, etwa wenn nun ein Konsens dahingehend bestünde, dass Risiken nicht mehr auf mehrere Schultern verteilt werden, sondern krankenversicherungstechnisch nur noch individuell berechnet werden sollen. Wenn jemand unbedingt Fallschirmspringen will, dann muss er eben 1.000 Euro Krankenversicherung im Monat bezahlen statt 200 Euro. Während dieses Beispiel gerade vor dem Hintergrund einer privaten Krankenversicherung noch einzuleuchten vermag, wird man

den Ansatz bei der gesetzlichen Krankenversicherung und der Verwirklichung des individuell eben nur bedingt steuerbaren Lebensrisikos Krankheit nicht umsetzen können, ohne mit dem Solidarprinzip in Konflikt zu geraten.

Die Frage ist, wie genau unsere Verfassung die Prinzipien Solidarität auf der einen Seite und Selbstverantwortlichkeit auf der anderen Seite austariert. Ausgangspunkt ist dabei das Sozialstaatsprinzip im Sinne des Artikels 20 unseres Grundgesetzes. Dabei handelt es sich aber um ein sehr ausgestaltungsbedürftiges Prinzip, d.h. dass es erst durch den Gesetzgeber klare Konturen erhält. Gesetzesänderungen, die auf die Abschaffung des Solidaritätsprinzips hinauslaufen, müsste man sich jedenfalls auch mit Blick auf die Ewigkeitsklausel und den änderungsfesten Kern unserer Verfassung sehr genau ansehen. 

4.3 Algorithmenethik

Weitere Konflikte im Umgang mit Big Data treten bezüglich der Algorithmen auf, die den Analysen und Auswertungen zugrunde liegen. Mit dieser Problematik befasst sich die bislang auf gesellschaftlicher und wissenschaftlicher Ebene nur in Ansätzen diskutierte Algorithmenethik, die oben bereits kurz angesprochen wurde.

Big-Data-Anwendungen beruhen auf Aussagen über Wahrscheinlichkeiten, die aufgrund bestimmter algorithmischer Rechenmodelle getroffen werden. Wenn eine Nutzerin weiblich, zwischen 30 und 40 Jahren alt ist und schon häufiger Bücher eines bestimmten Genres bei einem Online-Buchhändler erworben hat, dann lässt sich mit einer bestimmten Wahrscheinlichkeit voraussagen, dass sie an dem neu erscheinenden Buch eines bestimmten Autors Interesse zeigen wird. Für den Händler lohnt es sich dann, auf den Webseiten, die die Nutzerin besucht, Werbung für eben dieses Buch zu schalten.

Um solche Aussagen treffen zu können, konstruieren Algorithmen auf Grundlage der verfügbaren Datenbasis den Durchschnittsmenschen bzw. -konsumenten, dessen Präferenzen und Verhaltensmuster als „normal“ qualifiziert werden. Wenn Berechnungen einen dergestalt definierten Normalzustand zum Ausgangspunkt nehmen, dann kann das unter Umständen für solche Personen zum Problem werden, die zu weit von dieser Norm abweichen oder sonst in keine der davon abgeleiteten Kategorien fallen. Aus dieser Erwägung heraus erscheint es daher angezeigt, die Algorithmen anhand ethischer Maßstäbe zu überprüfen.

Ein fiktives Beispiel aus der Verkehrslenkung soll dies verdeutlichen. Das Tracken der Geodaten der Bürger einer Stadt kann für die Gemeinschaft vorteilhaft sein, wenn dadurch die Fahrpläne des öffentlichen Nahverkehrs optimiert werden können. Wenn allerdings eine bestimmte Bevölkerungsgruppe bei der Datenerhebung systematisch ignoriert wird, weil die Mitglieder der Gruppe aus ökonomischen oder anderen Gründen nur geringen Zugang zu Smartphones oder anderen Geräten haben, mit denen das Tracking erfolgt, dann kann dies dazu führen, dass das algorithmisch ermittelte Ergebnis der Datenauswertung an den Bedürfnissen dieser Gruppe vorbeigeht. Auf-

grund unvollständiger Datenlage konnte der Algorithmus diese Gruppe nicht erfassen und ist somit von vornherein diskriminierend.

Die Frage, die sich in diesem Zusammenhang stellt, ist also, wie gewährleistet werden kann, dass Algorithmen ethischen Maßstäben gerecht werden. Ein erster notwendiger Schritt wären Transparenzpflichten, die durch einen Digitalen Kodex festgelegt werden könnten. Denn nur so besteht überhaupt die Chance, die Berechnungsgrundlage von Big-Data-Anwendungen auf versteckte Diskriminierungen hin zu überprüfen. Daran anschließend wäre zu klären, was geschehen soll, wenn sich ein Algorithmus beziehungsweise die Datenbasis, die ihm zugrunde liegt, als unausgewogen herausstellt. Soll es eine Möglichkeit geben, den Algorithmus anzufechten, also eine Korrektur einzufordern? Wenn ja, wer soll das durchsetzen dürfen – soll es eine neutrale Instanz geben, die Algorithmen prüft und ihre Verwendung gegebenenfalls untersagt? Wie könnte ein solches Recht umgesetzt werden?

Weiterhin würde sich ein Digitaler Kodex für den Umgang mit Big Data anbieten, um die Frage zu erörtern, wie Regulierungsmöglichkeiten überprüft werden könnten. So könnte erwogen werden, der Markteinführung neuer Tracking- und sonstiger Big-Data-Anwendungen einen Überprüfungsmechanismus im Hinblick auf die eingesetzten Algorithmen vorzuschalten. Schlussendlich müsste in einem letzten Schritt geklärt werden, wie die Einhaltung des Kodex sichergestellt werden könnte und wer für die Überwachung der Einhaltung zuständig sein sollte.

„Wir brauchen eine Regulierung in Bezug auf die Nutzung von Daten. Allerdings ist es aufgrund des internationalen Charakters der Digitalisierung schwierig, diese nur auf nationaler Ebene anzusiedeln. Hier muss ein internationales Grundverständnis entstehen, auf dessen Basis Abkommen oder Regulierung fußen können.“

Sabrina Juran, Technical Specialist, Data and Research, United Nations Population Fund (UNFPA), Interview

5. Epilog

Der vorliegende Bericht schließt das Unterthema „Big Data“ im Rahmen der zweiten Phase des Projekts „Braucht Deutschland einen Digitalen Kodex?“ ab. Dennoch markiert er nur ein Zwischenergebnis. Ein Digitaler Kodex für den Umgang mit Big Data ist bewusst nicht formuliert worden. Bis zu diesem Punkt sollte es lediglich darum gehen, anhand der drei Beispiele die Chancen und Risiken von Tracking und Big Data zu identifizieren und zu diskutieren. Das Projekt hat gezeigt, dass eine Reihe von Konfliktlinien existiert, hinsichtlich derer Regulierungsbedarf besteht. Welche Möglichkeiten der Regulierung es in dem Feld gibt und welche Art der Regulierung am sinnvollsten ist, wäre in einem nächsten Schritt zu erörtern, bevor dann gegebenenfalls ein Kodex gemeinsam mit den relevanten Akteuren ausgehandelt werden kann.

Auf der Ebene der Europäischen Union wird zurzeit über die Kodifizierung eines neuen europäischen

Datenschutzrechts verhandelt. Auch hier sind die durch die Digitalisierung und die Verbreitung von Big Data aufgeworfenen Fragen ein Thema. Sollte die Datenschutzgrundverordnung noch 2015 verabschiedet werden, ist davon auszugehen, dass sie 2017 in Kraft tritt. Jüngste Entwicklungen im Juni 2015 deuten darauf hin, dass die EU-Staaten dazu tendieren, die traditionellen Datenschutzprinzipien der Zweckbindung und der Datensparsamkeit zugunsten der Möglichkeit kommerzieller Big-Data-Anwendungen ein Stück weit aufzuweichen. Hier bleibt abzuwarten, inwieweit bis zum endgültigen Beschluss eine Balance zwischen den widerstreitenden Interessen gefunden wird. Davon unabhängig wird die Grundverordnung kaum die Frage nach alternativen Regulierungsmöglichkeiten der Datenerhebung und -auswertung durch Tracking und Big Data obsolet werden lassen. Der Aushandlungsbedarf wird bestehen bleiben.

Annex

Im Annex findet sich ein Übersichtspapier zu Big Data im Volltext, das einen einführenden Überblick über den komplexen Themenbereich erlaubt. Dies ermöglicht dem Leser, bei Interesse die im Rahmen des Projekts gemachten Überlegungen im Detail nachzuvollziehen.

ANALYSE DES THEMENKOMPLEXES „BIG DATA“

1. Einleitung

Die Digitalisierung ermöglicht das Sammeln und Verarbeiten von Daten in einem neuen, bislang unbekanntem Ausmaß. Jede Äußerung – etwa wem ich wann eine E-Mail schreibe – und jede Aktivität – wann ich welche Webseite an welchem Ort anschau – kann gespeichert und ausgewertet werden. Durch die zunehmende „Verdatung“ der Welt – Sensoren in Haushaltsgeräten, Überwachungskameras im öffentlichen Raum, Funkzellenabfrage in der mobilen Telefonie – entstehen immer mehr Daten, die von verschiedenen Stellen ausgewertet und genutzt werden. Das sind auf der einen Seite der Staat und seine Organe, aber auch zunehmend private Unternehmen. Die bekanntesten

sind wohl die großen Internet-Unternehmen wie Facebook, Google und Amazon – sie sind aber nicht die einzigen.

Große Datensammlungen sind kein neues Phänomen. Von der Volkszählung 1987, die in Deutschland von einer kontroversen Debatte begleitet wurde, bis zum heutigen Internet der Dinge sind sie immer wieder von gesellschaftlichen Diskussionen begleitet worden. Wissenschaftler, Politiker, Kulturkritiker und Experten spekulieren über ihre gesellschaftlichen Auswirkungen – mit durchaus unterschiedlichen Ergebnissen.

Der Zeitpunkt und die begriffliche Ausrichtung der derzeitigen Debatte sind nicht zufällig. Die Analyse von umfangreichen Datenbeständen ist seit Jahrzehnten in Wirtschaft und Wissenschaft etabliert, jedoch haben neue Technologietrends zu einer erheblichen Beschleunigung geführt. In den letzten zehn Jahren hat sich sowohl die Verfügbar-

keit als auch die technische Machbarkeit von Analysen erheblich weiterentwickelt. Damit wird eine Vielzahl bislang unerschlossener Auswertungen möglich.

Mit Big Data lassen sich erhebliche wirtschaftliche Potenziale realisieren: Das können effektivere Werbeschaltungen auf Webseiten sein, wo den Nutzern anhand ihres Nutzungsprofils relevante Werbung angeboten wird, oder intelligente Thermostate, die ihre Heizleistung der tatsächlichen Nutzung anpassen, die aufgrund von Big-Data-Analysen errechnet wurde. Es entstehen neue Geschäftsmodelle und Produkte, die effizienter und intelligenter arbeiten – und im Idealfall Gewinn bringen.

In der Wissenschaft kann Big Data erkenntnisbringend eingesetzt werden: In der Medizin werden über Big-Data-Analysen Krankheitsursachen und Heilungsmethoden entwickelt, in der Soziologie können Bevölkerungsbewegungen analysiert und in der Linguistik neue automatische Übersetzungstechnologien eingeführt werden, um nur einige Einsatzmöglichkeiten zu nennen.

Big Data ist aber nicht nur positiv besetzt. Der Begriff schürt Ängste und Befürchtungen, insbesondere vor umfassenden Persönlichkeitsprofilen und Verhaltensprognosen. Viele Kritiker sehen einen umfassenden Überwachungsstaat auf uns zukommen. Wenn private Anbieter Daten sammeln und sie verarbeiten, erwerben sie intime Kenntnisse über die einzelnen Nutzer. Sie können daraus Rückschlüsse über das Individuum, einzelne Personengruppen oder die gesamte Gesellschaft ziehen. Problematisch daran ist, dass meist nicht gesichert ist, wer was mit den Daten macht, ob sie weitergegeben und mit anderen Daten in Beziehung gesetzt werden und wer welche Erkenntnisse aus ihnen ableitet. Die Einzelnen wissen in aller Regel nicht, welche Daten zu welchen Zwecken gesammelt werden, was mit ihnen geschieht und welche Chancen und Risiken die Sammlung und Verarbeitung für sie birgt. Selbst wenn Nutzer in einer Datenschutzerklärung der Verarbeitung zustimmen, treffen sie oft keine informierte Entscheidung, sondern akzeptieren einfach die Bedingungen, die vorge-

geben werden. Bequemlichkeit geht vor informationelle Selbstbestimmung.

Dazu kommt ein weiteres Problem: Selbst wenn ich zustimme, dass meine Daten ausgewertet werden dürfen, kann ich diese Entscheidung nicht für Dritte treffen. Genau dies passiert aber. Der New Yorker Rechtswissenschaftler Eben Moglen spitzte dies in einem Vortrag an der Columbia Law School zu, indem er sagte: „Privacy is not transactional.“¹ Sammlungen von persönlichen Daten wirken auch jenseits des Sammlers und des von der Sammlung Betroffenen. Denn die Verallgemeinerungen, die sich aus statistischen Auswertungen ergeben, führen zu Zuschreibungen auf Dritte. So funktioniert jedes Kredit scoring. Vereinfacht gesagt: Wenn bei 28-jährigen Männern Kredite häufiger ausfallen, bekommt ein 28-jähriger Mann keinen Kredit. Diese Drittbetroffenheit könnte gravierende Konsequenzen für bestehende Regulierungsansätze etwa des Datenschutzrechtes haben, die regelmäßig auf die Übereinkunft zwischen Datenverarbeiter und dem unmittelbar von der Verarbeitung Betroffenen setzen, Dritte aber nicht einbeziehen.

Der österreichische Jurist und Professor am Oxford Internet Institute Victor Mayer-Schönberger schlägt vor, sich darauf zu konzentrieren, wie Daten verarbeitet werden und wozu sie genutzt werden, statt ausschließlich auf die Frage, ob sie überhaupt erhoben werden.² Vieles spricht dafür, die Bearbeitungsprozesse stärker ins Blickfeld zu rücken. Neue Methoden der Datenanalyse haben auch in der Vergangenheit dazu geführt, dass gesellschaftliche Übereinkünfte neu geregelt werden müssen. Der Digitalunternehmer Stephan Noller forderte auf einer Sommerakademie des Unabhängigen Landesentrums für Datenschutz Schleswig-Holstein sogar die Einführung von ethischen Grundsätzen für Algorithmen, eine „Algorithmen-Ethik“.

Solche Aushandlungsprozesse sollten auf einem soliden Wissensfundament und einer ethischen Grundlage stattfinden. Damit ist das Thema Big Data auf mehreren Ebenen prädestiniert, im Rahmen des Projektes „Braucht Deutschland einen Digitalen Kodex?“ näher betrachtet zu werden.

1 Eben Moglen, „Snowden and the future, Part III: The Union, May it be Preserved“, Vortrag an der Columbia Law School, 13. November 2013. Online unter <http://snowdenandthefuture.info/PartIII.html>.

2 Viktor Mayer-Schönberger und Kenneth Cukier, „Big Data: Die Revolution, die unser Leben verändern wird“, 2013.

Unter dem Begriff „Big Data“ wird eine Vielzahl von – teilweise komplexen – technischen und gesellschaftlichen Verfahren aus unterschiedlichsten Anwendungsfeldern diskutiert und problematisiert. Um diese vielfältigen Sachverhalte greifbarer zu machen, umreißt dieser Text zunächst einige Beispiele, die veranschaulichen sollen, um welche Art der Datenverarbeitung es sich handelt und welche gesellschaftlichen Diskurse daraus folgen. Ausgehend von diesen Beispielen, identifizieren wir anschließend einige Grundlagen und Technologietrends, die für Big Data charakteristisch sind. Im Folgenden werden dann einige der Problemfelder und die Potenziale von Big Data verdeutlicht, um dann schließlich in der abschließenden Zusammenfassung in konkrete Fragestellungen für Aushandlungsprozesse zu münden, die einen Beitrag für einen weiteren Diskussionsprozess darstellen sollen.

2. Was ist Big Data? – exemplarische Nutzungsszenarien

Wie der Name schon sagt: Bei Big Data geht es um große Datenmengen. Die im Rahmen der Digitalisierung verfügbaren einzelnen Daten sind immens: Immer mehr Sensoren sammeln Daten. Das können Wetterdaten sein, Verkehrsdaten, Gen-Daten, große Sprachkorpora, Logdateien von Webservern und Nutzerinteraktionen, Überwachungskameras im öffentlichen Raum, RFID-Reader und so weiter. Big Data ist als Begriff weit gefasst und nicht fest definiert, sodass der Gebrauch des Begriffes weitere Aspekte umfassen kann.

Besonders wichtig für Big Data sind die neuen Möglichkeiten der Verarbeitung und Analyse. Durch parallelisierte Datenverarbeitung in großen Rechenzentren lassen sich diese Datenmengen teilweise in Echtzeit verarbeiten. Dabei ändert sich die Herangehensweise: Man sucht nicht mehr in den Daten nach Beweisen für vorher entworfene Theorien, sondern untersucht die Datensätze nach Korrelationen und lässt sie selbst „sprechen“, wie es Viktor Mayer-Schönberger in seinem Buch ausdrückt.

Weitgehend akzeptiert ist die Analyse großer Datenmengen wohl im Rahmen klassischer, akademischer Forschung. Sie funktioniert regelmäßig einwilligungsbasiert und ist sowohl in Hinblick auf den Zweck, die Verfahren und Ziele weitgehend transparent. Zudem werden nach akademischer Tradition die Ergebnisse veröffentlicht und dienen damit der Allgemeinheit.³ Grundrechtlich schafft die Wissenschaftsfreiheit des Grundgesetzes eine befürwortende Wertungslage.

Ebenfalls weitgehend unumstritten sind jene Bereiche, in denen nicht die Beobachtung von Personen Gegenstand der Erhebung ist, etwa beim Large Hadron Collider in Genf, der Sternen- und der Wetterbeobachtung. Problematisch und diskursrelevant scheinen damit vor allen Dingen jene großen Datensammlungen zu sein, die menschliche Verhaltensweisen zum Gegenstand haben. Dies gilt insbesondere dann, wenn sie jenseits wissenschaftlicher Tätigkeit in der Privatwirtschaft und im öffentlichen Sektor stattfinden.

Erhebung und Speicherung von großen Datenbeständen sind vielfach von gesellschaftlichen Aushandlungsprozessen begleitet. Diese können sich um das Ob der Erhebung, aber auch um das Wie der Verarbeitung drehen. Im Folgenden sollen exemplarisch verschiedene Bereiche nachgezeichnet werden, in denen Big Data eine Rolle spielt und die in der Gesellschaft mehr oder weniger breit diskutiert werden.

2.1. Der erste Big-Data-Diskurs in Deutschland: Das Volkszählungsurteil

Der Datenbestand der Volkszählung von 1987 nimmt mit einem Speichervolumen von geschätzt 80 Megabytes einen Umfang ein, der allenfalls Festplatten aus den 1980er-Jahren an ihre Grenzen bringt. Dennoch wird man bei diesem Datenbestand von Big Data sprechen können. Das Ziel war, aktuelle Zahlen über Bevölkerung, Versorgung, Verkehrsmittel und so weiter zu erhalten, um notwendige infrastrukturelle Maßnahmen einzuleiten.

Die Volkszählung ist ein gutes Beispiel dafür, wie die Erhebung von großen Datenbeständen öffentliche

³ Dennoch beginnt jüngst auch hier eine Debatte, vgl. die Gründung eines „Council for Big Data, Ethics, and Society“ in den USA, Meldung online unter: <http://www.whitehouse.gov/sites/default/files/microsites/ostp/Data2Action%20Announcements.pdf>.

Diskurse auslöst. Die für 1983 festgesetzte Volkszählung musste ausgesetzt werden und wurde schließlich im gleichen Jahr vom Bundesverfassungsgericht untersagt. Dieses Urteil des Bundesverfassungsgerichts setzte gewichtige Eckpfeiler für unser heutiges Datenschutzrecht, indem es den Grundsatz der „informationellen Selbstbestimmung“ definierte. Es machte zum einen Vorgaben für die Erhebung von Daten und etablierte zum anderen Prinzipien wie Datensicherheit und Zweckbindung der Daten, die für die anschließende Verarbeitung relevant sind.

2.2. Internet-Nutzungsdaten: Rezeptions- und Kommunikationsverhalten

Das jährliche Internet-Nutzungsverhalten einer Person inklusive des Inhalts aller besuchten Internet-Seiten beträgt bei einer durchschnittlichen Nutzung für drei Stunden pro Tag etwa 65 Gigabytes und könnte gut lokal auf einer handelsüblichen Festplatte abgelegt werden.

Der US-Mathematiker und -Unternehmer Stephen Wolfram protokolliert seit vielen Jahren die Interaktion mit seinen Rechnern. Zusätzlich zu der Analyse seiner E-Mail-Daten loggt er mit einem sogenannten Keylogger jede seiner Tastatureingaben mit. Wie er zeigt, lassen sich daraus interessante Ableitungen gewinnen. Er kann ablesen, welche Anwendungen er in den vergangenen Jahren intensiv genutzt hat und dass seine Fehlerraten beim Tippen relativ konstant geblieben sind.⁴ Hierbei entstehen auch bei intensiver Nutzung pro Tag Datenmengen im Bereich von lediglich 100 Kilobytes. Wolfram reiht sich mit seinen Selbstprotokollen in eine Bewegung ein, die unter dem Begriff „Quantified Self“ bekannt geworden ist – das „quantifizierte Ich“. Dabei tracken die Anwender etwa Gesundheitsdaten, wie oft sie Sport machen, zählen ihre Schritte oder andere persönliche Daten. Erste Geschäftsideen und Applikationen, die eine solche Selbstbeobachtung unterstützen, existieren bereits: etwa Fitnessarmbänder, die aufzeichnen, wie oft man sich bewegt, ob man ausreichend schläft und

Ähnliches. Bei der neuen iOS-Version von Apple ist die „Health“-App schon vorinstalliert, mit der Nutzer ihre Blutwerte, Gewichtsdaten oder Vitalwerte aufzeichnen können. Solange diese Daten beim Nutzer selbst bleiben und er Auswertungen seines Nutzungsverhaltens selbst vornimmt, hat dies keine problematischen Auswirkungen. Zum Problem wird es erst, wenn diese Daten zentral gesammelt und von externen Firmen ausgewertet werden. Auch anonymisiert können daraus Profile erzeugt werden.

Für die Online-Werbevermarktung sind das interessante Daten. Große Teile des globalen Nutzungsverhaltens werden bereits an zentralen Stellen erhoben und ausgewertet. Dieses ist begleitet von einer kontroversen öffentlichen Debatte. So gab es kürzlich eine intensive Diskussion um die Einführung einer „Do Not Track“-Funktion für Browser, die es den Nutzern erlauben sollte, selbst zu steuern, welche Webseiten ihn tracken dürfen und welche nicht. Die damit verbundenen Implikationen für bestehende Geschäftsmodelle haben zu erheblichen Kontroversen in technischen Standardisierungsgremien geführt. Auch der deutsche und europäische Gesetzgeber hat sich wiederholt kontrovers mit der Regulierung der Online-Werbung befasst und die Grundlagen für eine Erhebung differenziert (und kompliziert) ausgestaltet.⁵

Neben dem Tracking des Rezeptionsverhaltens entstehen im Internet umfassende Datenbestände über das Kommunikationsverhalten. Diese Datenbestände bestehen auf der einen Seite aus den eigentlichen Inhalten der Kommunikation (zum Beispiel dem Wortlaut von E-Mails, Nachrichten oder Chats), die im Rahmen gezielter Werbeschaltung analysiert werden, auf der anderen vor allen Dingen aus den sogenannten Metadaten der Kommunikation, also wer wann mit wem kommuniziert hat. Diese Daten sind – insbesondere unter Anwendung von Big-Data-Methoden – ausgesprochen aufschlussreich.

Ein Beispiel dafür ist die Analyse der Kommunikationsmetadaten von Arbeitnehmern, aus der sich Kündigungswahrscheinlichkeiten ableiten lassen.⁶ Der

⁴ Vgl. <http://blog.stephenwolfram.com/2012/03/the-personal-analytics-of-my-life/>.

⁵ Insbesondere in den Absätzen 3, 3a und 3b des § 28 Bundesdatenschutzgesetz (BDSG).

⁶ Vgl. etwa Frank Rieger: „Der Mensch wird zum Datensatz“. Veröffentlicht unter anderem auf FAZ.NET, dem Online-Auftritt der Frankfurter Allgemeinen Zeitung, 16. Januar 2010, online: <http://www.faz.net/aktuell/feuilleton/ein-echtzeit-experiment-der-mensch-wird-zum-datensatz-1591336.html>.

Arbeitgeber kann so schon vor dem Arbeitnehmer abschätzen, ob er dem Betrieb erhalten bleibt oder eigentlich schon auf dem Absprung ist. Ob eine solche Analyse in Deutschland rechtlich zulässig ist, ist zweifelhaft, aber nicht in allen Fällen eindeutig. Eine Überarbeitung des Arbeitnehmerdatenschutzes ist seit Jahren auch in Deutschland in der Diskussion, aber ein entsprechender Kabinettsentwurf wurde Anfang 2013 wieder von der Agenda genommen. Sollten sich solche Verfahren etablieren, ist wohl mit einer erheblichen Dynamisierung des bisher eher schleppenden parlamentarischen Beratungsprozesses bei der Novellierung des Arbeitnehmerdatenschutzes zu rechnen.

2.3. Sammlung von Nutzerdaten durch Geheimdienste

Neben der Sammlung von Nutzerdaten von Seiten privater Firmen hat seit 2013 – mit den Enthüllungen von Edward Snowden – die Datensammlung durch Geheimdienste, insbesondere durch die NSA, für eine erhebliche öffentliche Auseinandersetzung gesorgt.

Über die Speicherkapazitäten der NSA existieren unterschiedliche und bisweilen absurd erscheinende Spekulationen. Selbst wenn man ersten Einschätzungen der Magazine „Der Spiegel“⁷ und „Wired“⁸ nicht trauen mag, die Kapazitäten seien im Bereich von „Yottabytes“ (ein Yottabyte sind 10^{24} Bytes oder etwa das 1000-Fache der von der Firma Cisco antizipierten globalen Internet-Kommunikation für 2015⁹), erscheint es nicht ganz abwegig, dass eine Speicherung aller digital vermittelten menschlichen Kommunikations- und Rezeptionssachverhalte in Kürze möglich sein wird. Eine detaillierte Analyse, was und wie viel die Geheimdienste sammeln, steht trotz des großen öffentlichen Interesses noch aus. Ein sachlicher öffentlicher Diskurs ist jedoch schwierig, weil es um staatliche Geheimnisse geht. Der gesellschaftliche Aushandlungsprozess findet trotzdem statt, wenngleich unter erschwerten Bedingungen und in einer frühen Phase.

2.4. Sensordaten: das Internet der Dinge

Ebenfalls noch nicht abgeschlossen sein dürften die Aushandlungsprozesse über die Gefahren und Potenziale des sogenannten Internets der Dinge, also etwa moderne Hausmesstechnik und -steuerung (Smart Home, Smart Metering, Smart Grid), Smart Cars (bis hin zu selbst fahrenden Fahrzeugen) und Smart Watches (die als Fitnessstracker und PDAs dienen).

Bei der privaten Heizungstechnik, besonders bei Geräten mit starkem Elektrizitätsverbrauch, sind derzeit erhebliche digitale Produktinnovationen zu beobachten. Über die in diesen Produkten enthaltene Messtechnik werden detaillierte Daten erhoben: wann die Bewohner da sind, welches ihre bevorzugte Raumtemperatur ist und vieles mehr. Die darauf basierenden Innovationen versprechen erhebliche Einsparungen für den Energieverbrauch.

Im vergangenen Jahr hat die Markteinführung von „Nest“, einem „smarten“ Thermostat, öffentliche Aufmerksamkeit erfahren. Das Thermostat sammelt neben den Angaben über den Installationsort Sensordaten und Nutzungsdaten des Geräts und sendet sie an das Unternehmen Nest. Laut Datenschutzerklärung misst das Thermostat Raumtemperatur, Luftfeuchtigkeit und Lichteinwirkung. Welche Sensoren das Gerät hierfür verwendet, teilt die Firma nicht mit, genauso wenig, wie häufig gemessen wird. Darüber hinaus meldet ein Bewegungsmelder, ob sich etwas im Raum bewegt.¹⁰ Das Nest-Thermostat lernt aus dem Verhalten der Nutzer, die das Gerät über ihr Smartphone steuern können. Auf diesem Wege soll der Verbrauch von Heizung und Klimatechnik um bis zu 26 Prozent reduziert werden können. Anfang 2014 wurde das Unternehmen von Google zum Preis von 3,2 Milliarden US-Dollar übernommen. Allerdings ist bisher weitgehend unbekannt, wie Nest die entstehenden Daten verarbeitet.

Die Verbreitung von verbesserter Mess- und Steuerungstechnik bei der Elektrizität durch „intelligente Zähler“ soll nicht nur den Stromverbrauch in privaten

⁷ <http://www.spiegel.de/netzwelt/netzpolitik/bluffdale-das-datensammel-zentrum-der-nsa-a-904355.html>.

⁸ http://www.wired.com/2012/03/ff_nsadatacenter/all/.

⁹ Cisco (Hg.), „Cisco Visual Networking Index: Forecast and Methodology, 2013–2018“, online: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.pdf.

¹⁰ Vgl. <http://nest.com/legal/privacy-statement/>.

Haushalten senken, sondern auch Lastunterschiede im Stromnetz besser nivellieren können.¹¹ Weitgehend unbemerkt von der Öffentlichkeit hat hierzu ein umfangreicher Fachdiskurs stattgefunden, der zur Entwicklung zahlreicher technischer Normen geführt hat, die den Datenschutz- und IT-Sicherheitsaspekten einen Stellenwert zumessen.¹² Neben wirtschaftswissenschaftlichen Untersuchungen wird die Einführung intelligenter Stromzähler in Deutschland auch durch Forschung über die datenschutzrechtlichen Implikationen begleitet.¹³ Ihr Gegenstand sind die Mechanismen, die eine Datenerhebung und Weitergabe im Sinne der intendierten Analyseverfahren sicherstellen sollen. Der Diskurs nimmt hier also sowohl die Prozesse als auch die Verfahren ins Blickfeld.

Fitness-Tracker – Geräte zur Messung der eigenen, insbesondere körperlichen, Aktivität – finden in den letzten Jahren zunehmend einen Markt. Die Geräte erfassen neben Schritten und zurückgelegter Strecke zumeist die Pulsfrequenz. Krankenversicherungen beginnen nun, Interesse an den entstehenden Datenbeständen zu entwickeln. Je nach gemessener sportlicher Betätigung gewähren einzelne Versicherungen Prämien, so zum Beispiel der Versicherer Generali.¹⁴

Diesem Produktbereich kommt eine besondere Relevanz zu, weil völlig neue, bisher unerschlossene Datenquellen zugänglich gemacht werden, die gleichzeitig einen hochgradig persönlichen Charakter aufweisen und Rückschlüsse auf den sensiblen Bereich der Gesundheitsdaten ermöglichen. Gleichzeitig steht eine Vielzahl neuer Produkte in den Startlöchern, so zum Beispiel die „Apple Watch“, mit der man die Zeit ablesen, telefonieren, SMS schreiben, aber auch seinen Herzschlag und sein Bewegungsprofil aufzeichnen kann. Die Mitbewerber von Samsung, Sony und LG sind schon mit Android-basierten Smartwatches auf dem Markt, Motorola und Apple folgten 2015. Wel-

che Anwendungen möglich sind, ist noch nicht absehbar. Über Potenziale und Risiken muss noch diskutiert werden. Ebenso wenig ist absehbar, welche neuen Verfahren zur Auswertung der neuen Datenmassen entstehen werden.

Neben Smartwatches und Fitness-Trackern beschäftigten sich die Medien in letzter Zeit vermehrt mit Geräten zur Protokollierung des Fahrverhaltens von Autofahrern. Auch hier stellt sich die Frage, ob die umfassenden Datensammlungen, die dadurch entstehen, nicht zu sehr in das Persönlichkeitsrecht eingreifen. Die Geräte können detaillierte Bewegungsprofile herstellen, zeichnen Geschwindigkeit und Fahrverhalten auf, optimieren den Kraftstoffverbrauch und so weiter. Einige Versicherungen erwägen, auf der Grundlage solcher Daten teilnehmenden Autofahrern abhängig von ihrem Fahrverhalten günstigere Versicherungsprämien einzuräumen. Den datenschutzrechtlichen Bedenken wollen die Hersteller durch Privacy-by-Design begegnen. Durch technische Modellierung der Verfahren soll verhindert werden, dass detailliertere persönliche Daten entstehen, als für die Auswertung nötig sind. Die Daten sind aber auch für andere Analysen interessant: Die Diskussion über geschlechtsspezifische Unterschiede im Fahrverhalten könnten auf der Basis konkreter Daten geführt werden, eine Verknüpfung mit der Pulsmessung durch Fitness-Tracker könnte Hinweise geben, wann der Fahrer oder die Fahrerin sich eine Pause gönnen sollte oder Ähnliches.

3. Trends und Grundlagen

Big Data ist vielschichtig, es gibt keine einheitliche, feste Definition für diesen Begriff. Um ihn und die Implikationen zu (be)greifen, ist es notwendig, die mit

11 Hauptsächlich auf indirekte Effekte abstellend und ohne konkrete Zahlen zu den tatsächlich erwarteten Einsparungen beim Energieverbrauch: <http://www.bmwi.de/BMWi/Redaktion/PDF/Publikationen/Studien/kosten-nutzen-analyse-fuer-flaechendeckenden-einsatz-intelligenterzaehler,property=pdf,bereich=bmwi2012,sprache=de,rwb=true.pdf>. Für den Bereich privater Haushalte ist die Sinnhaftigkeit des Einsatzes unterdessen allerdings wohl fraglich, vgl. <https://www.bdew.de/internet.nsf/id/eine-frage-der-perspektive-de>, dort unter Verweis auf eine im Auftrag der RWE in Mülheim durchgeführte Breitenstudie, nach der nur relativ geringe Effizienzsteigerungen möglich sein sollen.

12 Bundesamt für Sicherheit in der Informationstechnologie (Hg.), „Schutzprofil für die Kommunikationseinheit eines intelligenten Messsystems für Stoff- und Energiemengen [BSI-CC-PP-0073]“, online: https://www.bsi.bund.de/DE/Themen/DigitaleGesellschaft/SmartMeter/smartmeter_node.html, sowie International Standardisation Organisation (Hg.), „ISO/IEC TR 27019:2013 Information technology – Security techniques – Information security management guidelines based on ISO/IEC 27002 for process control systems specific to the energy utility industry“.

13 Oliver Raabe et al., „Datenschutz in Smart Grids: Anmerkungen und Anregungen“, 2011.

14 Ralf Grötter, „Der Gläserne Patient“, Handelsblatt Online, 26.11.2014, online: <http://www.handelsblatt.com/technologie/forschung-medizin/medizin/der-glaeserne-patient-du-musst-dein-leben-aendern/11030186.html>.

dem Begriff verknüpften technischen Trends im Blick zu haben. Neben der reinen Datenmenge bezieht sich der Begriff typischerweise auf die „drei V“: Volumen, Velocity (Geschwindigkeit) und Varietät, der hier um die Eigenschaft der Verfügbarkeit ergänzt wird. Schließlich spielt regelmäßig Wahrscheinlichkeitsrechnung in verschiedenen Formen, zum Beispiel als selbstlernende Verfahren („Machine Learning“), eine große Rolle.

Datenanalysen kommen in vielen Anwendungsfeldern zum Einsatz. Nicht minder vielfältig sind die Verfahren, die dabei verwendet werden: „Query and Reporting“¹⁵, „Data-Mining“¹⁶, Datenvisualisierung, Vorhersagemodelle und Prognosen, Optimierung (vor allem von Prozessen), Simulation, Integration verschiedener Datenformate (etwa bei Sprach- und Bilderkennung), Geodaten- und raumbezogene Analysen sind nur einige.

3.1. „Volume“ und Parallelisierung

Grundsätzlich sind auch sehr große Datenmengen mit aktuellen technischen Architekturen in überschaubarer Zeit verarbeitbar. Es stehen inzwischen alternative Ansätze zu traditionellen Datenbankmodellen zur Verfügung, die es ermöglichen, die Verarbeitung von Daten auf parallel arbeitenden Computern zu koordinieren. Damit können Datenmengen jenseits von mehreren Gigabytes verarbeitet werden – zu erschwinglichen Kosten.

Der Suchindex von Google – weltweit die meistgenutzte Suchmaschine – hat derzeit einen Umfang von 100 Petabytes (oder 100 Millionen Gigabytes).¹⁷ Er wird kontinuierlich fortgeschrieben, indem immer neue Webseiten indiziert und gespeichert werden. Die Milliarden täglichen Suchanfragen¹⁸ zu beantworten, kann nur funktionieren, wenn sie gleichzeitig auf parallel geschalteten Computern bearbeitet werden. Diese Parallelisierung ist weder technisch trivial, noch war sie immer selbstverständlich. Sie hat erst in den letz-

ten Jahren an Popularität gewonnen und ist inzwischen im Rahmen von Cloud-basierten Infrastrukturen leicht nutzbar.

3.2. „Velocity“ – Just-in-time-Verarbeitung

Die Zeit, die für die Verarbeitung großer Datenbestände notwendig ist, stellt nach wie vor einen limitierenden Faktor dar, weil ein Ergebnis nur dann hilfreich ist, wenn es sehr schnell verfügbar ist („just in time“). Es kommt also auf den Einzelfall an, ob bestimmte Verarbeitungen möglich und sinnvoll sind.

Soll neben der Beantwortung einer Suchanfrage gleichzeitig die passende Werbung eingeblendet werden und auf das aktuelle Surfverhalten abgestimmt sein, stellt dies hohe Anforderungen an die Geschwindigkeit der Verarbeitung. Erfolgt die Auswahl der richtigen Werbung nicht rechtzeitig für das Suchergebnis, kommt sie zu spät.¹⁹ Verzögert sich die Beantwortung der Suchanfrage jenseits gewisser hinnehmbarer Grenzen, besteht die Gefahr, dass die Suchmaschine Marktanteile verliert. Mit der Verbreitung von Real-Time-Bidding-Systemen, die es ermöglichen, automatisiert im Augenblick der getätigten Suchanfrage für die Platzierung einer Anzeige ein Gebot abzugeben, gibt es einen erheblichen Bedarf für Systeme, die den Wert eines Gebots schnell auf einer großen Datenbasis errechnen können.

3.3. „Variety“ – Unstrukturierte Daten

Eine weitere Herausforderung für die Verarbeitung stellt nach wie vor die Struktur der vorhandenen Daten dar. Einzelne Datenkomponenten in Datenbanken sind stets durch einen Feldtyp näher beschrieben, beispielsweise Vorname, Nachname, Geburtsdatum.

Den Großteil aller vorhandenen Datenbestände wird man als unstrukturiert bezeichnen können. Ein Prosatext beispielsweise weist in der Regel keine annähernd exakten Strukturen auf; dazu ist die

¹⁵ Deutsch: „Suchen und Berichten“.

¹⁶ Der Begriff wird uneinheitlich verwendet. Man kann hierunter das gezielte Auffinden einzelner Informationen in größeren Datenbeständen verstehen, wobei hier analytische Verfahren zum Einsatz kommen können, die bestehende Informationen verknüpfen.

¹⁷ Google publiziert hierzu nicht regelmäßig. Die genannte Zahl wird in 2010 und 2012 erwähnt: <http://googleblog.blogspot.de/2010/06/our-new-search-index-caffeine.html> und hier: <http://googleforwork.blogspot.de/2012/07/introducing-google-cloud-platform.html>.

¹⁸ Derzeit sind es ca. 3,5 Milliarden pro Tag, vgl. <http://www.internetlivestats.com/google-search-statistics/>.

¹⁹ Anschaulich illustriert in dem Video von Matt Cutts, Software-Ingenieur bei Google, „How Search Works“, <https://www.youtube.com/watch?v=BNHR6lQJZs>.

Grammatik und Semantik natürlicher Sprachen zu komplex, flexibel und mehrdeutig. Schwierigkeiten für die Analyse können entstehen, wenn Daten aus unterschiedlichen Quellen mit zwar vorhandenen, aber nicht übereinstimmenden Strukturmerkmalen auftauchen. Die Übergänge sind jedoch fließend, weshalb die Verwendung des Begriffs „unstrukturierte Daten“ gelegentlich als unpräzise bezeichnet wird. Für den Umgang mit unstrukturierten Daten existiert eine Vielzahl von Verfahren, von denen einige im weiteren Verlauf exemplarisch näher beleuchtet werden sollen.

3.4. Verfügbarkeitsanforderungen als Treiber für Innovation

Die Verfügbarkeit von Datenbeständen spielt für die möglichen Verarbeitungen eine zentrale Rolle. Diese Frage gerät in den typischen kommerziellen Einsatzszenarien leicht aus dem Blickfeld, da Unternehmen meist vom eigenen Datenbestand ausgehen und gegebenenfalls zukaufbare Datenbestände heranziehen und nach Auswertungswegen und Erkenntnismöglichkeiten suchen, wenn sie überlegen, wo ihnen Big Data helfen kann.

Es ist aber auch der umgekehrte Weg denkbar, indem man sich der Frage nähert, in welchen Bereichen man Erkenntnisse aus Datenbeständen gewinnen möchte und anschließend nach Möglichkeiten sucht, diese Datenbestände aufzubauen.

Neben Start-ups, deren Geschäftsmodelle direkt auf Datenanalyse begründet sind, ist dieser Blickwinkel auch für große Unternehmen relevant, die ihre strategische Geschäftsentwicklung und Akquise zunehmend daran ausrichten, welche Datenbestände hierdurch erschließbar werden. Die starken Bewertungen und Kaufpreise von Firmen mit solchen Datenbeständen oder Potenzialen in diesen Bereichen sprechen eine deutliche Sprache. Beispiele hierfür sind offenkundig die Unternehmen Facebook und Google selbst, aber auch die schon genannte Firma

Nest. Nicht abwegig ist wohl die Überlegung, dass Google sich nach der Entwicklung seiner Kernkompetenz im Bereich der Internet-Suche gezielt auf weitere Anwendungsfelder wie E-Mail, Soziale Netzwerke und Kartografie bewegt hat, um damit Zugang zu Datenbeständen zu gewinnen, die mit den bestehenden korreliert werden können. Über die genauen Verknüpfungen dieser Datenbestände und die angewendeten Verfahren ist indes wenig bekannt. Dies fällt unter den Schutz der Geschäftsgeheimnisse des Unternehmens.²⁰

Innovationstreiber ist demnach zunehmend, aus der Big-Data-Analyse Erkenntnisse zu gewinnen, die dann als Geschäftsgeheimnis einen mitunter relevanten Teil des Marktwertes des Unternehmens ausmachen. Es stellt sich allerdings die Frage, ob dies ein guter Innovationsanreiz ist oder ob er zu intransparenten und mitunter sogar marktverzerrenden Geschäftsmodellen führt.

3.5. Wahrscheinlichkeiten und Machine Learning

Eine wichtige Grundlage zum Verständnis von Datenanalysen ist in vielen Fällen die Wahrscheinlichkeitsrechnung. Nahezu allen Verfahren ist eine gewisse Verwurzelung im Methodenkanon der Stochastik gemein, wobei der Wahrscheinlichkeitsrechnung besondere Bedeutung zukommt.

In manchen Bereichen sind Wahrscheinlichkeits-erwägungen allerdings nur am Rande relevant oder spielen gar keine Rolle. Zu denken ist etwa an Datenvisualisierungen oder auch konventionelle Anfragen, in deren Rahmen es nur darum geht, eine bestimmte Information in einem großen Datenbestand auffindbar zu machen („Query-Verfahren“).²¹ Allerdings können hier auch stochastische Methoden integriert werden. Für die Suche nach einzelnen Informationen in unstrukturierten Datenbeständen kann man über Wahrscheinlichkeiten etwa die Anzahl der näher zu untersuchenden Datenbestände minimieren.

²⁰ Insbesondere im Fall des schon 2007 erfolgten Zukaufs des Werbedienstleisters DoubleClick ist die Nähe zum Kerngeschäft offensichtlich, eine Verknüpfung der Datenbestände liegt nahe. Google hat zwar versichert, dass eine Verknüpfung auf der Ebene der personenbezogenen Profile im Rahmen der Übernahme nicht erfolgt. Das heißt aber nicht unbedingt, dass abstrahierte Datensätze, die wegen einer erfolgten Aggregation keinen Personenbezug aufweisen, nicht miteinander kombiniert wurden.

²¹ Derartige Suchen nach der sprichwörtlichen Nadel im Heuhaufen könnte man begrifflich auch dem Data-Mining zuordnen. Tatsächlich werden die Begriffe Data-Mining und Big Data beide nicht streng definiert und nicht selten nahezu synonym verwendet.

Ein Grundprinzip aller Wahrscheinlichkeitsangaben ist, dass sie keine Aussagen über den Einzelfall zulassen, sondern nur Prognosen im Rahmen von (bei korrekter Anwendung klar definierten) Wahrscheinlichkeiten ermöglichen. Dabei ist der Begriff der Korrelation zentral. Eine Korrelation beschreibt die Abhängigkeit einer Größe von einer anderen. Ist die Korrelation hoch, kann man (wiederum im Rahmen der Wahrscheinlichkeiten) vom einen Wert eine Aussage für den anderen ableiten.

Ein klassisches Beispiel für einfache Datenanalysen ist die Prozessoptimierung. In wiederkehrenden Prozessen können Daten über Produktionsfehler erfasst werden. Findet sich in den Daten eine Korrelation der Produktionsfehler mit der Tageszeit, ergibt das einen Anhaltspunkt für eine Optimierung. Um ein ganz einfaches Beispiel zu nennen: So kann eine Korrelation darauf hindeuten, dass Mitarbeiter am Fließband im Idealfall nach drei Stunden eine Pause machen müssen oder dass ein Industrieroboter nach 48 Stunden gewartet und mit neuem Schmieröl versorgt werden muss. Korrelationen über die Zeit spielen auch in der Verkehrsplanung eine Rolle. Belastungssituationen ergeben sich zu bestimmten Stoßzeiten. Um den Verkehrsfluss zu optimieren, werden Ampelschaltungen und Fahrpläne, aber auch die Verkehrsplanung und der Infrastrukturausbau auf der Basis von Verkehrszählungen verbessert.

Zunehmend werden bei der Datenanalyse auch selbstlernende Verfahren eingesetzt, die aus der Künstliche-Intelligenz-Forschung stammen. Mustererkennungsverfahren oder maschinelles Lernen mittels neuronaler Netze oder vereinfachter Ansätze ermöglichen das Auffinden von regelhaften Strukturen in Datenbeständen. Auch diese Verfahren fußen auf statistischen Berechnungen.

Ein beliebtes Beispiel für den Einsatz von Big-Data-Analysen sind automatisierte Übersetzungen. Für diesen Bereich haben sich lernende Mechanismen den Ansätzen als überlegen gezeigt, deren Regeln vorher determiniert sind. Nahezu alle automatischen Übersetzungsdienste bauen darauf auf, durch die Analyse eines größeren Textkorpus Regeln zu gewin-

nen, die sie dann auf neue Texte, deren Übersetzung nicht bekannt ist, anwenden. Das Faszinierende an diesen Verfahren ist, dass es für das System weitgehend egal ist, für welche Sprachen es trainiert wird. Entscheidender ist die Masse (und Qualität) des Datenbestandes, aus dem „gelernt“ wird. Frühere Ansätze, die auf bestehenden grammatikalischen Regeln und Wörterbüchern aufgebaut waren, sind weniger erfolgreich gewesen. Ähnliche Verfahren werden nun zur Optimierung der Spracherkennung eingesetzt.²²

4. Potenziale und Herausforderungen

Big Data bietet enorme Potenziale, weist aber auch erheblichen gesellschaftlichen und – daraus folgend – politischen und rechtlichen Klärungsbedarf auf. Bestimmend für die wachsende Bedeutung von Datenanalysen ist die der Digitalisierung innewohnende Verdichtung nahezu aller Lebensbereiche.

Sowohl Potenziale als auch Probleme stellen sich den verschiedenen Akteursgruppen mit jeweils spezifischen Handlungsoptionen. Die Potenziale für die einzelnen Akteure sind vielfältig, ebenso wie die Herausforderungen, aus denen im Nachfolgenden fünf zentrale herausgehoben werden sollen.

Durch die Verfügbarkeit umfassender Datenbestände ergeben sich so beispielsweise für staatliche Akteure neue Möglichkeiten, ihre Entscheidungen präziser an den gesellschaftlichen Bedürfnissen auszurichten. Diese Potenziale werden bisher nur zögerlich genutzt. In der Stadtplanung wird zwar in nahezu jedem Politikbereich auf Empirie zurückgegriffen. Die Frage ist aber, ob durch eine stärkere Verdichtung nicht ganz neue Qualitäten politischer Gestaltung in einer Vielzahl von Feldern erreicht werden können.

Die Nutzung von Big Data gehört in vielen Wissenschaftsdisziplinen zum grundlegenden Handwerk. Wissenschaftler sind – wenn man so will – die natürlichen Early Adopter. Allerdings ist die Bandbreite erheblich: Während in der experimentellen Elementarteilchenphysik, wie sie am Genfer Institut für Elementarteilchenphysik CERN praktiziert

²² Zum Einsatz von neuronalen Netzen bei der Spracherkennung siehe etwa: <http://www.forbes.com/sites/roberthof/2013/05/01/meet-the-guy-who-helped-google-beat-apples-siri/>.

wird, Big Data zum Kerngeschäft gehört, ist in der Rechtswissenschaft die Datenanalyse jenseits der Kriminalistik bisher kaum ein Thema. Es ist aber davon auszugehen, dass auch in datenfernen Disziplinen über Datenanalysen interessante Erkenntnisse gewonnen werden können. Entsprechend werden in vielen Bereichen neue Wege beschritten. So versucht das vom Bildungsministerium geförderte Projekt „Argumentum“ eine automatisierte Analyse von Argumentationsstrukturen in Gerichtsurteilen zu entwickeln.²³ Aber schon eine – weitaus einfachere – Analyse von Gesetzesverweisen in Urteilen und anderen Rechtstexten könnte Aufschluss darüber geben, welche Normen praktisch relevant sind, und so effektiv zu einer Verschlankung des Rechtsbestandes führen.

Erhebliche Potenziale bieten sich in der empirischen Sozialforschung. Dieses Gebiet, dessen Primärdaten bisher in nicht unerheblichem Maße aus Umfragen gespeist waren, kann unter Zugriff auf Beobachtungsdaten, wie sie im Internet-Zeitalter entstehen, völlig neuen Wirkzusammenhängen nachgehen. Zentrales Problem in diesem wie auch in anliegenden Feldern wie der Psychologie und sicherlich auch der Medizin ist jedoch die Verfügbarkeit des Datenmaterials für wissenschaftliche Forschungszwecke.

Die Möglichkeiten der wirtschaftlichen Nutzung von Big Data werden zunehmend erschlossen. Nach einer repräsentativen Studie des Branchenverbands Bitkom werten neun von zehn Unternehmen grundsätzlich IT-gestützt Daten für ihre Entscheidungsprozesse aus; 46 Prozent der Unternehmen setzten 2013 dafür bereits spezielle Analyse-Tools ein; weitere 36 Prozent beabsichtigten, 2014 Maßnahmen in diesem Bereich einzuführen.²⁴ Nachteile ergeben sich mitunter für kleinere Unternehmen, die nicht über geeignete Datenbestände verfügen. Große Anbieter haben daher eine bessere Ausgangsposition für Optimierungen ihrer Geschäftsmodelle.

Hier sind allerdings in einigen Bereichen auch Disruptionen bestehender Märkte denkbar. So ver-

fügen Landwirte beispielsweise nicht über hinreichende Informationen über ihre Endkunden, da sie ihre Produkte meist über Zwischenhändler vertreiben. Diese Zwischenhändler profitieren von ihrem besseren Marktwissen und können so mitunter erhebliche Preisaufschläge rechtfertigen. Denkbar ist, dass sich Landwirte – zumindest in einigen spezialisierten Märkten – über Vertriebsplattformen zusammenschließen, um diese Informationen zu sammeln und den Vertrieb weitgehend selbst zu organisieren. Die Datenorganisation, die früher eine Domäne des Vertriebs war, fällt damit zurück in die Hände der Erzeuger.²⁵

Auch aus individueller Perspektive ist Big Data zunehmend ein Thema. Die bereits erwähnte „Quantified Self“-Bewegung, die sich der quantifizierbaren Selbsterfassung verschrieben hat, findet zunehmend in die Alltagstechnologie Einzug. Über Pulsmessgeräte mit digitaler Schnittstelle und Schrittzähler wird für Hobbysportler eine Trainingsoptimierung möglich, wie sie früher nur Leistungssportlern zugänglich war.

GPS-Systeme in Smartphones machen eine detaillierte Auswertung der eigenen Wege möglich. Es ist davon auszugehen, dass zukünftig eine Vielzahl von Informationen des Alltags sensorisch erfasst wird und damit detailliert Aufschluss etwa über Ernährungsgewohnheiten, Gesundheitsentwicklung oder über Lernfortschritte in der Schule geben kann.

Viele dieser Geräte sind jedoch an spezialisierte Dienste gebunden und übermitteln die erfassten Daten umfassend an den Hersteller oder Diensteanbieter. Hierdurch wird es dem Nutzer erleichtert, den Austausch und Vergleich mit anderen zu suchen. Welche Daten übermittelt werden, ist allerdings nicht immer transparent. Bei manchen Daten, zum Beispiel gesundheitsrelevanten, ist eine Übermittlung aufgrund ihrer Sensitivität nicht wünschenswert. Darüber hinaus bestehen Probleme bezüglich der Profilbildung und Manipulation.

23 Vgl. <http://argumentum.eear.eu>.

24 Vgl. Bitkom (Hg.): „Potenziale und Einsatz von Big Data – Ergebnisse einer repräsentativen Befragung von Unternehmen in Deutschland“, Berlin, 5.5.2014, online: http://www.bitkom.org/files/documents/Studienbericht_Big_Data_in_deutschen_Unternehmen.pdf, wobei sämtliche Angaben sich auf Unternehmen mit einer Größe von mehr als 50 Mitarbeitern beschränken.

25 Diesen Weg geht etwa das kanadische Start-up „Provender“, vgl. <http://www.theglobeandmail.com/report-on-business/small-business/sb-growth/day-to-day/farms-factories-and-film-sets-startups-bring-big-data-to-inefficient-industries/article20574000/>.

4.1. Erkenntnisgewinn, aber: Korrelation ist nicht Kausalität

Eine gute Datenanalyse birgt die Chance auf Erkenntnisse, die sich für Wirtschaft, Wissenschaft und politische Gestaltung nutzbar machen lassen. Doch Korrelation bedeutet nicht Kausalität: Nur weil man Abhängigkeiten zwischen zwei Größen erkennt, heißt das nicht, dass die eine die andere verursacht. Es kann auch genau umgekehrt sein, oder beide Größen hängen gleichermaßen von einer dritten ab. Und schließlich kann es auch reiner Zufall sein: Die Größen sind unabhängig voneinander. Das klassische Beispiel hierfür ist die Korrelation zwischen dem Bestand an Störchen und der Geburtenrate in einem Dorf.²⁶ Selbst wenn jedes Mal, wenn ein Storch ins Dorf kommt, auch ein Kind geboren wird, heißt das noch nicht zwingend, dass die Störche die Kinder bringen.

Allerdings gilt auch: ohne Korrelation keine Kausalität. Kommt ein Kind zur Welt, ohne dass ein Storch im Dorf war, ist damit jedenfalls belegt, dass nicht (nur) Störche Kinder bringen, sondern Kinder auch ohne Störche zur Welt kommen können. Eine nicht ganz unwichtige Erkenntnis.

Die Korrelation bietet durchaus eine gute Grundlage für die Suche nach tatsächlichen Sachzusammenhängen, insbesondere aber zur Falsifikation von angenommenen Sachzusammenhängen.

Das Problem aktueller Big-Data-Analysen ist jedoch, dass sie so viele verschiedene Faktoren gleichzeitig in eine Korrelationsanalyse einbeziehen, dass einzelne, mögliche Ursachenzusammenhänge nicht mehr erkennbar sind. Ergebnis: Man weiß, dass unter bestimmten Bedingungen mit einer hohen Wahrscheinlichkeit ein bestimmtes Ereignis eintritt, hat aber keine Ahnung, wieso.

4.2. Vermeintliche Vollerhebungen und blinde Flecken

War man vor der Digitalisierung auf Umfragen und Zählungen angewiesen, um Datenmaterial für Analysen zu erhalten, entstehen heute diese Datenbestän-

de in ungleich höherer Quantität und zumeist auch Qualität quasi nebenbei.

Ein relevanter Teil traditioneller empirischer Marktforschung erfolgt auf der Basis von Umfragen. Auf diese Weise versucht man zum Beispiel, die Effektivität einer Werbemaßnahme zu erfassen und gegebenenfalls anzupassen. Die gesammelten Aussagen sind allerdings nur dann aussagekräftig, wenn es gelingt, eine repräsentative Auswahl von Befragten zu bestimmen und zur Teilnahme zu bewegen. Selbst dann bleiben die Aussagen mit einer gewissen Ungenauigkeit behaftet. Verzerrungen ergeben sich dadurch, dass Befragte im Einzelfall absichtlich oder unabsichtlich falsche Angaben machen. Bezieht jedoch ein Datenbestand nicht nur einen Teil, sondern alle (potenziellen) Kunden ein und baut er zudem auf der direkten Verhaltensbeobachtung auf, ist er von erheblich höherer Qualität. Dies geschieht beispielweise beim Online-Marketing. Hierbei wird das Nutzungsverhalten detailliert aufgezeichnet. Firmen, die auf Online-Marketing spezialisiert sind, können zuordnen, welche Internet-Seiten ein Nutzer besucht hat. Die Spuren, die ein Nutzer hinterlässt, geben Aufschluss darüber, wer welchen Werbemitteln ausgesetzt war und wo diese in eine Kaufentscheidung münden.

Eine gute Datenanalyse muss sich aber der blinden Flecken in den Datenbeständen gegenwärtig sein. Prognoseaussagen können immer nur bezogen auf die Grundgesamtheit getroffen werden. Eine Prognose, die auf der Verhaltensbeobachtung von Nutzern bestimmter Sozialer Netzwerke aufbaut, ist – selbst wenn das Soziale Netzwerk erhebliche Größe hat – nicht auf den Rest einer Gesellschaft übertragbar. Große Datenbestände laden mitunter zu unsachgemäßen Verallgemeinerungen ein. Es ist sehr verlockend, auf die Datenbestände Verfahren anzuwenden, die eigentlich nicht zu richtigen Ergebnissen führen. Gerade aufgrund der Komplexität der Verfahren sind solche Fehler nicht immer leicht nachzuvollziehen. Baut man Entscheidungen gesamtgesellschaftlicher Bedeutung auf derartigen Analysen auf, läuft man Gefahr, großen Bevölkerungsgruppen nicht gerecht zu werden.

²⁶ Siehe z.B. „Storch und Mensch“, http://www.zeit.de/2006/25/Stimmt-s_P-25_xml.

4.3. Exklusives Wissen und Marktversagen

Erkenntnisgewinn grundlegender Art war bislang die Domäne der Wissenschaft. Aufgrund ihres Charakters orientiert sich Wissenschaft an dem Grundsatz, diese Erkenntnisse nicht exklusiv zu behandeln, sondern zu publizieren. Dies ändert sich, wenn die Grundlagen für den Erkenntnisgewinn exklusiv in privater Hand sind.

Für Marktakteure bieten Datenanalysen eine Vielzahl von Optimierungsmöglichkeiten. Gleichzeitig können sie – sofern die Daten nur exklusiv verfügbar sind – zu erheblichen Verzerrungen führen und wünschenswerte Disruptionen und Innovationen aufhalten. Die Markteintrittsbarriere für Suchmaschinen ist nicht deswegen so hoch, weil es technisch aufwendig ist, den Bestand an Internet-Seiten zu erfassen und zu indizieren, sondern weil es an der Datenbasis fehlt, die Seiten danach zu gewichten, welche besonders stark frequentiert werden. Die Überlegung mag in verschärfter Weise für die Entwicklung eines nicht nur inhaltlich guten, sondern auch rentablen Geschäftsmodells gelten, da der Markt bereits stark auf – wiederum exklusive – Datenbestände der Werbevermarktung optimiert ist.

4.4. Umfassende Profilbildung

Für Individuen bieten die genannten Möglichkeiten oft positive Effekte. Nutzer profitieren von Erkenntnissen und Produktoptimierungen, weil die Produkte dadurch preiswerter angeboten werden können. Jedoch gibt es auch Gefahren: Offenkundig ist das Problem der Profilbildung, die sich aus der umfassenden Digitalisierung ergibt. Datensammler können ihre Datenbestände, etwa über das Surfverhalten, auf einen individuellen Nutzer zurückführen. Über Korrelationen können sie dann Wahrscheinlichkeitsaussagen zu Fragen treffen, die den Nutzer in sensiblen Bereichen betreffen, ohne dass er dazu jemals direkt Auskunft gegeben hat. Es ist so-

gar denkbar, dass aus den Daten auf Informationen geschlossen wird, die darüber hinausgehen, was das Individuum über sich selbst weiß, etwa wenn auf Grundlage des Verhaltens in Sozialen Netzwerken Aussagen über das Vorliegen von unerkannten Krankheiten oder von sexuellen Vorlieben extrapoliert werden. Studien deuten darauf hin, dass im Einzelfall durchaus Aussagen von erheblicher Genauigkeit möglich sind. So wurde nachgewiesen, dass in Sozialen Netzwerken signifikante Aussagen über die sexuelle Orientierung einzelner Mitglieder gemacht werden können. Dabei wurden die Kontakte der Mitglieder ausgewertet (eine sogenannte Netzwerkanalyse). Auch wenn das betreffende Mitglied keine Auskunft über seine eigene sexuelle Orientierung abgegeben hat, konnte, gestützt auf die Angaben der jeweiligen Kontakte, ein Rückschluss darauf gezogen werden.²⁷

Eine ähnliche Untersuchung gibt es auch zur Depressionsneigung. Hierbei wurde das Kommunikationsverhalten von Personen, deren Depressionserkrankung bekannt ist, mit dem von Menschen korreliert, bei denen die Krankheit nicht diagnostiziert worden war. Wenn bestimmte Charakteristiken, etwa die Verringerung sozialer Aktivität, eine negative Grundstimmung in den Inhalten der Kommunikation oder stärkere Religiosität auftauchen, lieferte dies Anhaltspunkte für die Depressionswahrscheinlichkeit.²⁸ Die Untersuchung zeigt, dass eine Aussagegenauigkeit von etwa 70 Prozent möglich war, wenn die Eigenschaften an Personen getestet wurden, bei denen die Erkrankung bekannt war, aber diese Tatsache erst nach der Überprüfung berücksichtigt wurde.

Hinzu tritt eine weitere Gefahr: Aussagen sind nicht nur über die Mitglieder eines Sozialen Netzwerkes möglich, sondern auch über Dritte, die nicht Mitglied sind oder je waren. Eine Studie von Wissenschaftlern an der ETH Zürich zeigt, dass Netzwerkanalysen über sexuelle Vorlieben auch Aussagen über Externe zulassen.²⁹

27 Carter Jernigan, Behram F.T. Mistree, „Gaydar: Facebook friendships expose sexual orientation“, *First Monday*, 14.10 [2009], online: <http://firstmonday.org/article/view/2611/2302>.

28 Munmun De Choudhury et al., „Predicting Depression via Social Media“, *ICWSM*, 2013, online: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/viewFile/6124/6351>.

29 Emre Sarigöl, David Garcia, Frank Schweitzer, „Online Privacy as a Collective Phenomenon“, *CoRR*, arXiv:1409.6197, online: <http://arxiv.org/abs/1409.6197>.

4.5. Manipulation der Meinungsbildung

Die Gefahr der Manipulation auf der Basis von Verhaltensprognosemöglichkeiten, die einzelnen Organisationen exklusiv vorliegen, betrifft nicht nur das Individuum. Der Präsidentschaftswahlkampf von Barack Obama im Jahr 2012 war stark auf die Analyse von Daten Sozialer Netzwerke in den entscheidenden Swing-States gestützt. Es spricht einiges dafür, dass diese Herangehensweise einen entscheidenden Beitrag zu seinem späteren (knappen) Wahlerfolg geleistet hat. Aufgrund tief gehender Datenanalysen konnte sich das Team von Barack Obama effektiv auf besonders relevante Wählerschichten konzentrieren und die Themen identifizieren, mit denen diese Gruppen überzeugt werden konnten.³⁰

Steht derartiges Wissen nur einzelnen Kandidaten zur Verfügung, gefährdet das die Chancengleichheit verschiedener politischer Gruppen. Das wiederum gefährdet die Demokratie und damit die Grundlage des gesellschaftlichen Zusammenhalts. Haben die Bürger Grund für den Verdacht, dass Meinungsbildung nicht frei stattfindet, sondern gezielt gesteuert wird, schwindet das Vertrauen in die Demokratie. Damit droht ihre Bereitschaft zu sinken, Einschränkungen individueller Freiheit und Vorlieben zugunsten einer (aus ihrer Sicht bis dato noch unterstellten) gesellschaftlichen Mehrheit zu akzeptieren.

5. Zu diskutierende Fragestellungen

Überall dort, wo Probleme und Interessenkonflikte zutage treten, besteht die Notwendigkeit von Aushandlungsprozessen. Diese können entweder auf der Basis konkreter Szenarien geschehen, die dann abstrahiert auf ähnliche Fragestellungen angewendet werden, oder die Gesellschaft einigt sich auf abstrakter Ebene über die Grundsätze, die wichtig genug sind, um in die Aushandlungsprozesse einzufließen.

Da aber die Einsatzgebiete von Big Data so unterschiedlich sind, ist es sinnvoll, im Augenblick auf der Ebene der konkreten Einsatzgebiete zu bleiben. Es lassen sich aber zumindest strukturelle Überle-

gungen anstellen, die aufzeigen, wie Aushandlungsprozesse gestaltet werden können. Derzeit herrschen allerdings noch grundlegende Aufgabenstellungen vor, die Räume für individuelle Aushandlungsprozesse stark einschränken. Hier bedarf es zunächst einer grundsätzlichen gesellschaftlichen Einigung.

Im Rahmen des Projekts „Braucht Deutschland einen Digitalen Kodex?“ ergeben sich somit mehrere mögliche Ansatzpunkte. Dabei empfiehlt sich eine Verknüpfung abstrakter und konkreter Fragen. Letztere sind notwendig, weil – wie dargestellt wurde – sich viele Fragen nur und erst im konkreten Einsatzszenario stellen. Dennoch ist zu vermuten, dass sich auch auf der Metaebene grundsätzlichere Erkenntnisse gewinnen lassen. Hierbei könnte es hilfreich sein, neben der grundsätzlichen Diskussion auch einem „Bottom-up“-Zugriff Raum zu gewähren und aus der Diskussion einzelner Big-Data-Einsatzszenarien Abstraktionen zu ermöglichen.

Bezüglich der konkreten Einsatzszenarien, die zu erörtern sind, sind diejenigen vielversprechend, die bisher nicht Gegenstand einer umfassenden Diskussion waren, weil die technischen und markt betreffenden Entwicklungen noch recht jung sind, also etwa der Bereich der Smart Watches und Selbsttracker. Daneben könnte man erwägen, die Möglichkeiten umfassender Datensammlungen bei staatlichen Diensten oder den großen wirtschaftlichen Akteuren ins Blickfeld zu nehmen. Eine ganze Reihe gesellschaftlicher Herausforderungen stellen sich nur bei diesen umfassenden Sammlungen, die insoweit eine eigene Qualität haben.

Für die Diskussion von verallgemeinerbaren Überlegungen sollen die folgenden Abschnitte einen ersten Ansatzpunkt liefern. Sie stellen die rechtlichen Instrumentarien dar inklusive der Beschränktheit derzeitiger Regelungsansätze im Rahmen des Datenschutzes (Abschnitt 5.1) und beleuchten schließlich allgemeine, prozedurale und steuernde Aspekte für einzelne Verfahren (Abschnitt 5.2). Diese beiden Abschnitte beleuchten das Problem gleichsam aus zwei Perspektiven, zunächst „top down“ und dann „bottom up“. Es ist davon auszugehen, dass über die Betrachtung einzelner Verfahren abstrakte Verallge-

³⁰ Sasha Issenberg, „How President Obama’s campaign used big data to rally individual voters“, MIT Technology Review, 19.12.2012, online: <http://www.technologyreview.com/featuredstory/509026/how-obamas-team-used-big-data-to-rally-voters/>.

meinerungen gefunden werden können. Das Projekt könnte diese Erwägungen zum Ausgangspunkt für eine Diskussion nehmen und im Laufe des Prozesses weiterentwickeln.

Neben der Frage nach einem Digitalen Kodex für den privaten Bereich stellt sich naturgemäß auch die Frage der Regulierung bei staatlicher Nutzung von Big Data. Dieser Bereich wird jedoch für den folgenden Prozess zunächst ausgeklammert.

5.1. Selbstregulierung durch den Markt?

In einem Rechtsstaat wie Deutschland, der auf allgemeiner Handlungsfreiheit und Privatautonomie aufgebaut ist, ist bei Interessenkonflikten zunächst einer direkten Aushandlung zwischen den beteiligten Parteien der Vorzug zu geben. Staatliche Eingriffe in diese Aushandlungen gehen stets mit einem Eingriff in die Handlungsfreiheit einher und bedürfen daher einer Rechtfertigung.

Dieses Primat privater Aushandlungsprozesse kann sich nicht nur auf diese grundrechtlichen Erwägungen stützen. Die oben beschriebene Vielzahl von einzelnen Anwendungsfällen würde zu einer extrem kleinteiligen Regulierung führen, die aus Sicht der rechtsschöpfenden Politik nicht oder nur mit hohem Aufwand kompetent zu bewerkstelligen ist.

Staatliches Handeln ist jedoch dort geboten, wo die privaten Aushandlungsprozesse versagen, etwa weil sich ungleiche Parteien begegnen oder weil schutzwürdige Belange Dritter bzw. der Allgemeinheit betroffen sind.

Das Datenschutzrecht trägt der allgemeinen Handlungsfreiheit – insoweit es um den Schutz der informationellen Selbstbestimmung geht – dadurch Rechnung, dass es Datenverarbeitungen personenbezogener Daten dann gestattet, wenn diese für Vertragszwecke der beteiligten Parteien erforderlich sind. Dort, wo es nicht um Verträge geht, können die von der Verarbeitung Betroffenen explizit einwilligen. Das Ungleichgewicht zwischen unter Umständen sehr großen Diensteanbietern und Konsumenten berücksichtigt das Datenschutzrecht dadurch, dass es den Diensteanbietern Transparenzverpflichtungen auferlegt, etwa dass sie die Daten nur für die Zwecke verarbeiten dürfen, die beide Parteien miteinander vereinbaren.

Big-Data-Ansätze zeigen aber deutlich die Beschränktheit dieses Modells auf, das auf den übereinstimmenden Willen der beteiligten Parteien abstellt. Wie schon dargelegt wurde, sind von der Verarbeitung gegebenenfalls Personen betroffen, die ihre Daten gar nicht oder nur in geringerem Umfang zur Verfügung gestellt haben. So gelangt der eingangs zitierte Eben Moglen zu der Erkenntnis, die Privatsphäre als „nicht transaktional“ zu charakterisieren, was heißt, dass sie keine abgeschlossene Austauschbeziehung zwischen zwei Parteien ist. In Fällen der Betroffenheit Dritter scheitert die Privatautonomie. Hier besteht ein Problem, an das sich der Gesetzgeber in Deutschland bisher noch nicht herangewagt hat.

Hinzu tritt hier ein weiteres Problem: Der Anwendungsbereich des traditionellen Datenschutzrechts bezieht sich auf die Verarbeitung von personenbezogenen Daten. In der Datenbasis muss aber Personenbezug im Sinne herkömmlicher Definitionen gar nicht bestehen, um statistisch untermauerte Aussagen über Menschengruppen mit bestimmten Eigenschaften zu machen. Allenfalls wird er im Augenblick der Aussage offenkundig, was das Recht wohl nach bisher vorherrschender Ansicht nicht erfasst.

Verschiedentlich wurde gezeigt, dass asymmetrische Machtverhältnisse zwischen einzelnen Parteien im Kontext von Big Data eine neue qualitative Dimension erreicht haben könnten. Soweit solche Wissensbestände zu Marktversagen führen, wird dies kartellrechtlich relevant. Vereinzelt ist bereits die Forderung nach stärkeren Offenlegungspflichten beispielsweise für Suchalgorithmen zu hören. Im staatlichen Bereich weisen Informationsfreiheitsgesetze in eine ähnliche Richtung.

Welche Folgen die Analysemöglichkeiten haben, die aus einer umfassenden Verdattung des Alltags resultieren, ist noch nicht abschätzbar. Insbesondere stellt sich die Frage nach gesamtgesellschaftlichen Steuerungserfordernissen und -möglichkeiten. Hierbei werden vielfach grundlegende Fragen aufgeworfen. Eine Klärung zentraler Herausforderungen, wie die der Drittbetroffenheit und neuer Machtasymmetrien, ist vordringlich. Auf dieser Basis können dann Modelle für kleinteiligere Aushandlungsprozesse entwickelt werden.

Problemlagen wie diesen kommt das Recht regelmäßig über staatliche Verbote mit Genehmigungsver-

fahren bei. Doch verbieten und genehmigen kann man nur, wenn man weiß, wann. Im ersteren Fall muss das mit der Handlung verbundene Risiko hinreichend hoch sein. Im zweiten Fall muss deutlich sein, dass im konkreten Fall das Risiko anders zu bewerten ist. Doch was sind die Risiken? Und wann sind sie handhabbar? Wir wissen es noch nicht – jedenfalls nicht in allen Fällen.

Hieraus ergeben sich folgende übergeordnete Fragen:

- A.1. Gibt es Fälle, in denen die Datenerhebung nicht (ausschließlich) auf eine Einwilligung des Betroffenen gestützt werden kann? Fraglich ist dies vor allem in Fällen, in denen aus diesen Daten Rückschlüsse auf Dritte, die nicht eingewilligt haben, möglich sind.

Wenn dem so ist:

- A.2. In welchen Fällen gilt das? Wo also bedarf es ergänzender Mechanismen? Wenn man auf den Begriff des Risikos abstellt, welche Risiken sind das konkret? Gibt es hierfür definierbare, abstrakte Risikoschwellen? Ist beispielsweise die Anzahl der Datensätze ein Kriterium?

- A.3. Welche Regulierungsinstrumente kommen hierfür in Betracht? Genehmigungsverfahren? Generelle Verbote? Weitere? Wie kommen ergänzende Regelungen zustande? Durch staatliche Maßnahmen oder alternative Regulierungsansätze?³¹

5.2. Eckpunkte für die Einordnung einzelner Verfahren

Die Etablierung von neuen Datenanalysen bringt auch historisch regelmäßig auf den Einzelfall bezogene gesellschaftliche Aushandlungsprozesse mit sich. Diese finden typischerweise zu zwei Zeitpunkten statt, die gelegentlich auch zusammenfallen. Eine erste Aushandlung findet dann statt, wenn deutlich wird, dass bestimmte Datenbestände anfallen und

damit der Auswertung zugänglich werden, ein weiterer dann, wenn deutlich wird, welche Auswertungen vorgenommen und welche Erkenntnisse daraus gewonnen werden können. Ein dritter Fall – vom zweiten in der Praxis freilich nur graduell zu unterscheidender Prozess – kann dann eintreten, wenn deutlich wird, welche Handlungen aus den gewonnenen Erkenntnissen folgen. Ausgehend von den obigen, abstrakten Fragestellungen, lassen sich auf der Ebene der konkreten Verfahren folgende Fragen aufwerfen und ableiten:

- B.1. Wann sind Aushandlungsprozesse für konkrete Analyseverfahren erforderlich? Welches sind die Indikatoren für deren Erforderlichkeit? Wer prüft, ob dies der Fall ist?

Bereits bekannt ist, dass die Analysen insbesondere dann problematisch werden, wenn Informationen über Personen gesammelt werden. Dass diese im Weiteren aber nicht als personenbezogene Daten im Sinne des Datenschutzrechts verarbeitet werden müssen, wurde gezeigt. Das Datenschutzrecht greift hier (zumindest derzeit) als Instrumentarium wohl zu kurz.

- B.2. Gibt es verallgemeinerbare prozedurale Elemente für derartige Aushandlungsprozesse? Welche Möglichkeiten gibt es, und wann sind sie einschlägig?

Zu denken ist etwa an

- den politisch-parlamentarischen Prozess,
- behördliches Ermessen,
- technische Standards,
- industrielle Selbstverpflichtungen,
- daneben aber auch an neue Multi-Stakeholder-Prozesse,

die hier Anregungen liefern können. Insbesondere die drei letztgenannten Punkte bieten dabei jedoch noch erheblichen Raum, prozedurale Fragen auch

³¹ Siehe hierzu auch im Folgenden 5.2, insbesondere dort B.2.

vor dem Hintergrund der konkreten Hausforderungen durch Big Data weiter zu erörtern.

- B.3. Welches sind die Aushandlungslinien? Was kann Gegenstand der Aushandlung sein? Mit anderen Worten: Welche Anforderungen sind an die Durchführung von Big-Data-Analysen im konkreten Einzelfall zu stellen?

Erkennbar ist hier etwa das Kriterium Transparenz, und zwar entlang der drei Achsen

- Erhebung (welche Daten werden erfasst?),
- Verarbeitung (in welchen Verfahren werden die Daten verarbeitet, welche Erkenntnisarten können daraus entstehen und welche Erkenntnisse werden gewonnen?) und
- Nutzung (der Erkenntnisse) inklusive der Rückwirkungen (also etwa Preisdiskriminierungen).

Entsprechend können für jede relevante Big-Data-Analyse im Sinne der Frage B.1. (siehe oben) die folgenden Fragen gestellt werden:

- B.4. Ist in dem konkreten Prozess die Erhebung der Daten offenzulegen? Ist sie etwa in einem öffentlichen Register zu hinterlegen?³²

- B.5. Ist der konkrete Verarbeitungsprozess öffentlich zu machen? Welche Erkenntnisse können gewonnen werden? Welcher Algorithmus wird verwendet? Wie ist dieser implementiert?

- B.6. Ist zu veröffentlichen, wie die gewonnenen Erkenntnisse tatsächlich genutzt werden?

- B.7. Gibt es Fälle, in denen die Erkenntnisse selbst zugänglich gemacht werden müssen?

Grundlegender, aber entlang derselben Achsen, sind auch Einschränkungen möglich: bestimmte Daten nicht einzubeziehen, bestimmte Verfahren nicht anzuwenden, bestimmte Erkenntnisse nicht wirken zu lassen. Ein Beispiel für Letzteres sind Antidiskriminierungsregeln.

Diese Kriterien stehen jedoch oft im Konflikt mit den Interessen der Verarbeiter, genau diese Informationen für sich zu behalten, etwa um Marktvorteile zu erzielen.

- B.8. Gibt es Verfahren, deren Anwendung schon grundsätzlich auszuschließen ist? Gibt es Erkenntnisse, die nicht gewonnen oder zumindest nicht verwertet werden sollten?

³² Öffentliche Register könnten in einem erheblichen Maß zur Transparenz beitragen. Zwar werden sie gerne als „Bürokratiemonster“ gegeißelt, aber in Zeiten, in denen Datenverarbeitung omnipräsent ist, erscheint es nicht einsichtig, warum die Führung von Registern einen erheblichen Aufwand mit sich bringen muss. Neben einer Übersicht für den individuell Betroffenen, wo Daten über ihn oder sie gespeichert sein könnten oder wer über Erkenntnisse verfügen könnte, eröffnet ein Register zudem den reizvollen Pfad von Big-Data-Analysen über Big Data.

Quellen- und Literaturhinweise

Biem, Alain u.a. (2010): „Real-Time Traffic Information Management using Stream Computing“, Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, S. 64 ff.

Bitkom (2015): „Big Data und Geschäftsmodell-Innovationen in der Praxis: 40+ Beispiele. Leitfaden“, <https://www.bitkom.org/Bitkom/Publikationen/Big-Data-und-Geschaeftsmodell-Innovationen-in-der-Praxis-40-Beispiele.html>

Dewri, Rinku u.a. (2013): „Inferring Trip Destinations From Driving Habits Data“, Colorado Research Institute for Security and Privacy, Department of Computer Science, University of Denver, <http://cs.du.edu/~rdewri/data/MyPapers/Conferences/2013WPES-Extended.pdf>

Forum Privatheit (Hg.) (2015): „Das versteckte Internet. Zu Hause – Im Auto – Am Körper“, White Paper, https://www.forum-privatheit.de/forum-privatheit-de/aktuelles/aktuelles_dokumente/White-Paper-2-Final_17.07.15-Druckversion.pdf

Geiselberger, Heinrich/Moorstedt, Tobias (Hg.) (2013): „Big Data: Das neue Versprechen der Allwissenheit“, Suhrkamp

Heureux, Alain (2015): „Regulatory Challenges for Big Data“, in T. Schwarz (Hg.): „Big Data im Marketing“, Haufe, S. 308 ff.

Hofstetter, Yvonne (2015): „Big Data und die Macht des Marktes“, 65 Aus Politik und Zeitgeschichte 11-12, S. 33 ff.

Jähn, Karl/Nagel, Eckhard (2014): „e-Health“, Springer

Kreibich, Rolf (2015): „Von Big zu Smart – zu Sustainable?“, 65 Aus Politik und Zeitgeschichte 11-12, S. 20 ff.

Lahmann, Henning (2015): „Gesellschaftliche Konfliktlinien im Kontext von Big Data am Beispiel von Smart Health und Smart Mobility“, DIVSI, https://www.divsi.de/wp-content/uploads/2015/09/Themenpapier_Konfliktlinien-Big_Data.pdf

Langkafel, Peter (2015): „Dr. Algorithmus? Big Data in der Medizin“, 65 Aus Politik und Zeitgeschichte 11-12, S. 27 ff.

Lupton, Deborah (2014): „Self-Tracking Modes: Reflexive Self-Monitoring and Data Practices“, Paper for the Conference „Imminent Citizenships: Personhood and Identity Politics in the Informatic Age“, Canberra, <http://ssrn.com/abstract=2483549>

- Lupton, Deborah (2016):** „Digital Health Technologies and Digital Data: New Ways of Monitoring, Measuring and Commodifying Human Embodiment, Health and Illness“, in Olleros u.a. (Hg.): „Research Handbook on Digital Transformations“, Edward Elgar, S. 6 ff.
- Mainzer, Klaus (2014):** „Die Berechnung der Welt. Von der Weltformel zu Big Data“, C.H. Beck
- Martin, Kirsten E. (2015):** „Ethical Issues in the Big Data Industry“, 14 MIS Quarterly Executive, S. 67 ff.
- Mayer-Schönberger, Viktor/Cukier, Kenneth (2013):** „Big Data: Die Revolution, die unser Leben verändern wird“, Redline
- Mayer-Schönberger, Viktor (2015):** „Zur Beschleunigung menschlicher Erkenntnis“, 65 Aus Politik und Zeitgeschichte 11-12, S. 14 ff.
- Morgenroth, Markus (2014):** „Sie kennen dich! Sie haben dich! Sie steuern dich!: Die wahre Macht der Datensammler“, Droemer HC
- Morozov, Evgeny (2015):** „Ich habe doch nichts zu verbergen“, 65 Aus Politik und Zeitgeschichte 11-12, S. 3 ff.
- Noller, Stephan (2012):** „Relevanz ist alles. Plädoyer für eine Algorithmen-Ethik“, Frankfurter Allgemeine Zeitung Online, 24. Oktober 2012, <http://www.faz.net/aktuell/feuilleton/debatten/plaedoyer-fuer-eine-algorithmen-ethik-relevanz-ist-alles-11934495.html>
- Odlum, Michelle/Yoon, Sunmoo (2015):** „What can we learn about Ebola outbreaks from tweets?“, 43 American Journal of Infection Control, S. 563 ff.
- Polonetsky, Jules/Tene, Omer (2013):** „Privacy and Big Data: Making Ends Meet“, 66 Stanford Law Review, S. 25 ff.
- Reichert, Ramón (2014):** „Big Data: Analysen zum digitalen Wandel von Wissen Macht und Ökonomie“, transcript
- Rothstein, Mark A. (2015):** „Ethical Issues in Big Data Health Research“, 43 Journal of Law, Medicine and Ethics, <http://ssrn.com/abstract=2535373>
- Schallaböck, Jan (2014):** „Big Data“, DIVSI, https://www.divsi.de/wp-content/uploads/2015/01/Themenpapier_Big-Data.pdf
- Selke, Stefan (2014):** „Lifeloggung. Wie die digitale Selbstvermessung unsere Gesellschaft verändert“, Econ
- Stegemann, Patrick (2013):** „Algorithmen sind keine Messer“, Bundeszentrale für politische Bildung Netzdebatte, 28. Oktober 2013, <https://www.bpb.de/dialog/netzdebatte/170865/algorithmen-sind-keine-messer>
- Stöcker, Christian (2015):** „Politikfeld Big Data“, 65 Aus Politik und Zeitgeschichte 11-12, S. 8 ff.
- Vossen, Gottfried (2015):** „Big Data: Daten sammeln, aggregieren, analysieren, nutzen“, in T. Schwarz (Hg.): „Big Data im Marketing“, Haufe, S. 35 ff.
- Will, Matthias Georg (2015):** „Privacy and Big Data: The Need for a Multi-Stakeholder Approach for Developing an Appropriate Privacy Regulation in the Age of Big Data“, Discussion Paper No. 2015-03, Halle/Wittenberg, <http://ssrn.com/abstract=2634970>

ÜBER DIE AN DIESEM BERICHT BETEILIGTEN ORGANISATIONEN UND INSTITUTIONEN

Deutsches Institut für Vertrauen und Sicherheit im Internet (DIVSI)

Das Deutsche Institut für Vertrauen und Sicherheit im Internet (DIVSI) hat das Ziel, einen offenen und transparenten Dialog über Vertrauen und Sicherheit im Netz zu organisieren. DIVSI sucht hierzu die gezielte Unterstützung von Wissenschaft und Forschung und arbeitet mit renommierten Instituten zusammen.

Das Internet hat zentrale Bedeutung für Gesellschaft, Wirtschaft und Kultur. Es revolutioniert unsere Arbeit und die Freizeit, unser Denken und die Kommunikation. Technologische Weiterentwicklungen und Netzwerke schaffen Raum für Ideen und offerieren vielfältige Möglichkeiten: von der Selbstentfaltung des Einzelnen über neuartige Lösungen und Geschäftsmodelle bis hin zur radikalen Veränderung etablierter Industrien und gewohnter Verhaltensweisen.

DIVSI möchte einen Beitrag zum Verständnis dieser hohen Bedeutung leisten, aber auch potenzielle Risiken im Umgang mit dem Internet untersuchen und analysieren. Aufklärungsarbeit soll für eine Sen-



sibilisierung und für eine Steigerung von Vertrauen und Sicherheit im Internet sorgen. DIVSI setzt für einen interdisziplinären Meinungs austausch zwischen Wissenschaft, Wirtschaft und Gesellschaft Impulse. Es bietet ein Forum für den Austausch ökonomischer, regulatorischer, rechtlicher, sozialer, kultureller und medienpolitischer Perspektiven. Untermauert wird dieses durch themenspezifische Tagungen und Veranstaltungen sowie strategische Projekte. Im Rahmen der Unterstützung von Wissenschaft und Forschung hat DIVSI der Technischen Universität München zu Jahresbeginn 2012 eine Professur für „Cyber Trust“ gestiftet.

Schirmherr von DIVSI ist Bundespräsident a. D. Prof. Dr. Roman Herzog. Als Vorsitzende des Beirats fungiert Prof. Dr. Claudia Eckert (Inhaberin des Lehrstuhls für IT-Sicherheit, TU München). Direktor des Deutschen Instituts für Vertrauen und Sicherheit im Internet mit Sitz in Hamburg ist Matthias Kammer. DIVSI, eine Initiative der Deutsche Post AG, arbeitet unabhängig und gemeinnützig.

iRights.Lab

Das iRights.Lab ist zu Beginn des Jahres 2012 als unabhängiger Think Tank zur Entwicklung von Strategien im Umgang mit den Veränderungen in der digitalen Welt gegründet worden. Damit erweitert das iRights.Lab das thematische Feld von iRights.info auf neue Bereiche. Dazu gehören angewandte Forschung, die Entwicklung von Strategien für Unternehmen, Politik und die öffentliche Hand, die praktische Gestaltung von Veränderungsprozessen und die Bereitstellung eines geschützten Raumes zum interdisziplinären fachlichen Austausch zwischen Experten.

Leitbild des iRights.Lab ist, die Möglichkeiten der Digitalisierung und des Netzes zum Vorteil der Öffent-

lichkeit und der Gesellschaft zu nutzen. Das iRights.Lab ist weder parteipolitisch noch an Unternehmen gebunden. Vielmehr werden Fragestellungen entwickelt und die möglichen Antworten erforscht – interdisziplinär, unabhängig, verständlich, ergebnisoffen. Welche rechtlichen Rahmenbedingungen gibt es, wie sehen die technischen Möglichkeiten aus, wie entwickeln sich politische Prozesse, wie verhält man sich auf neuen Märkten, wie kommuniziert man über soziale Netzwerke? In thematischen Labs bearbeitet das iRights.Lab Themen wie kreative Arbeit und Kreativwirtschaft, Innovation, Journalismus, kulturelles Erbe oder Open Content – sowohl inhaltlich als auch strategisch.



DIVSI Studien im Überblick



Das Recht auf Vergessenwerden (2015)

Die Entscheidung des EuGH zum „Recht auf Vergessenwerden“ lässt gleichwohl Fragen unbeantwortet, die im Spannungsfeld zwischen Persönlichkeitsrechten, Datenschutz und dem Recht auf Meinungs- und Pressefreiheit liegen. Dieser komplexen Problematik widmet sich diese Publikation und formuliert schließlich konkrete Empfehlungen für einen „Lösch-Kodex“.



DIVSI Studie Beteiligung im Internet – Wer beteiligt sich wie? (2015)

Was ist Beteiligung im Internet eigentlich genau? Wie und weshalb bringen Internet-Nutzer sich ein? Die zweite Studie im Rahmen des DIVSI Forschungsprogramms „Beteiligung im Netz“ untersucht Formen, Vorteile und Hürden der Beteiligung im Internet aus Sicht der DIVSI Internet-Milieus. In der qualitativen Untersuchung kommen dabei die Internetnutzer selbst zu Wort.



DIVSI U9-Studie: Kinder in der digitalen Welt (2015)

Wissenschaftlich untermauert wird aufgezeigt, wie Kinder zwischen 3 und 8 Jahren die digitale Welt für sich erschließen und welchen Einfluss die digitale Lebenswelt der Eltern auf die digitale Sozialisation der Kinder hat. Die DIVSI U9-Studie liefert eine Vielzahl von Fakten und stellt Ansätze zur Entwicklung von geeigneten Maßnahmen bereit.



DIVSI Studie – Daten: Ware und Währung (2014)

In einer repräsentativen Bevölkerungsbefragung untersucht DIVSI das Online-Nutzungs- und -Konsumverhalten in Deutschland. Im Fokus stehen Einstellungen der Internet-Nutzer zu Themen der Datensicherheit sowie Weiterverwendung von persönlichen Daten.



DIVSI Studie – Wissenswertes über den Umgang mit Smartphones (2014)

Über Smartphones sind Menschen heute nahezu ununterbrochen „online“. Mit steigendem Nutzungsumfang fällt dabei eine Vielzahl von Daten an. „Was geschieht mit meinen Daten?“ lautet daher die Leitfrage dieser Studie und nimmt die Möglichkeiten der Einsichtnahme und Einflussnahme durch Nutzer bei unterschiedlichen mobilen Betriebssystemen in den Blick.



Braucht Deutschland einen Digitalen Kodex? (2014)

Mit dem Projekt „Braucht Deutschland einen Digitalen Kodex?“ lotet DIVSI aus, ob ein Digitaler Kodex ein geeignetes Mittel ist, verbindliche Regeln im Internet auszuhandeln und durchzusetzen. Der Projektbericht steuert nicht nur zu diesem Gedanken Anregungen bei. Er bietet darüber hinaus generelle Anstöße, über die nachzudenken sicherlich lohnt.



DIVSI Studie zu Bereichen und Formen der Beteiligung im Internet (2014)

Das DIVSI Forschungsprogramm „Beteiligung im Netz“ leistet auf einer breiten theoretischen und empirischen Basis einen Beitrag zum öffentlichen Verständnis der Beteiligungschancen des Internets – und ihrer Voraussetzungen. Die Studie präsentiert einen ersten Schritt in diesem Vorhaben und verschafft einen Überblick über den heutigen Stand der Forschung.



DIVSI U25-Studie (2014)

Die DIVSI U25-Studie liefert erstmals fundierte Antworten auf Fragen, die das Verhalten der nachwachsenden Generation im Hinblick auf das Netz betreffen. Über die Nutzungsformen hinaus werden auch die Denk- und Handlungslogiken sowie der lebensweltliche Hintergrund untersucht.



DIVSI Studie zu Freiheit versus Regulierung im Internet (2013)

Wie sicher fühlen sich die Deutschen im Internet? Wie viel Freiheit und Selbstbestimmung wollen sie? Nach wie viel Regulierung wird verlangt? Die Studie zeigt ein detailliertes Bild des Nutzungsverhaltens der Deutschen im Internet und ihrer Wahrnehmung von Chancen und Risiken.



Entscheider-Studie zu Vertrauen und Sicherheit im Internet (2013)

Wie denken Entscheider über das Internet? Welchen Akteuren schreiben sie welche Verantwortung und welche Einflussmöglichkeiten zu? Was sagen sie zu Sicherheits- und Freiheitsbedürfnissen? Die Studie verdeutlicht erstmals, wie diejenigen über das Internet denken, die wesentlich die Spielregeln gestalten und Meinungsbilder prägen.



Meinungsführer-Studie „Wer gestaltet das Internet?“ (2012)

Wie gut kennen sich Meinungsführer im Netz aus? Wie schätzen sie ihre Einflussmöglichkeiten ein? Welche Chancen, Konfliktfelder und Risiken erwachsen daraus? In persönlichen Gesprächen wurden führende Repräsentanten aus Politik, Wirtschaft, Verwaltung, Wissenschaft und Verbänden interviewt.



Milieu-Studie zu Vertrauen und Sicherheit im Internet (2012) + Aktualisierung (2013)

Die Milieu-Studie differenziert erstmals unterschiedliche Zugangsweisen zum Thema Sicherheit und Datenschutz im Internet in Deutschland, basierend auf einer bevölkerungsrepräsentativen Typologie.

