

Big Data - eine neue Herausforderung für Datenschutz

aus: Vorgänge Nr. 200 (Heft 4/2012), S.74-82

Das Marketing der IT-Industrie zielt auf das Setzen von Schwerpunktthemen. Für 2013 steht offenkundig Big Data auf der Tagesordnung, im Vorjahr waren es etwa das Cloud Computing oder das Smart Metering. Der gezielt geschaffene Hype (verstanden auch als überzogenes Versprechen) um bestimmte Technologien soll die Einführung und Akzeptanz am Markt befeuern. Bei Big Data handelt es sich dabei um eine sog. Hope-, Hype- and Fear- Technologie[1], bei der zum Teil spekulative Erwartungen[2], aber zugleich auch Befürchtungen und die damit einhergehende Kommunikation interessierter Kreise den Blick auf das Phänomen eher vernebeln als erhellen und gleichwohl großen Einfluss auf Entscheidungen wie z. B. Forschungsförderung etc, ausüben können. Solche Technologiedebatten sind oft gekennzeichnet durch verkürzende Darstellungen von Ambivalenzen der Technik, Unklarheiten bei der Definition und Abgrenzung des Phänomens selbst, Technikdeterminismus und Intransparenz durch verdeckte Interessen.[3]

Vorratsdatenspeicherung, Drohnen, Biometrie, digitale Innenstadttotüberwachungen (INDECT, Real Time Crime Center Manhattan, New York), aber auch das sensorengestützte „Internet der Dinge“ u. a. mittels RfiD-Chips, Energiedaten aus Smart Grids und Cloud Computing sind Beispiele für sich weiter abzeichnende technische Interventionen, sowohl bei komplexen gesellschaftlichen Problemen wie z. B. der „Kriminalitätsprävention“ als auch schlicht zur Steigerung der Erträge von Unternehmen. Diese Anwendungen produzieren rein quantitativ betrachtet große Datenmengen (Big Data). Doch auch bereits laufende Infrastrukturen wie moderne Mobilfunknetze mit ihren Millionen Smartphonennutzern, Soziale Netzwerke und natürlich die Infrastruktur des Internet selbst produzieren allein schon gigantische Mengen maschinenerzeugter Kontextdaten für jeden Vorgang und inhaltlich unstrukturierte Einzeldaten, die zumeist in keine Kundendatenbank passen .[4] Das Gros dieser Daten unterliegt allein der Nutzung zu den jeweiligen primären Zwecken wie etwa der Bereitstellung bestimmter Dienste. Man ist deshalb bereits früh auf die Idee gekommen, zusätzlichen ökonomischen Nutzen aus den einmal vorliegenden Daten zu ziehen.

Was ist und wie neu ist Big Data?

In der Wahrnehmung des größten bundesdeutschen Verbandes der IT-Industrie BIT-KÖM definiert man Big Data so: „Big Data (als Trend) bezeichnet die wirtschaftlich sinnvolle Gewinnung und Nutzung entscheidungsrelevanter Erkenntnisse aus qualitativ vielfältigen und unterschiedlich strukturierten Informationen, die einem schnellen Wandel unterliegen und in bisher ungekanntem Umfang anfallen.“[5] Der Begriff Big Data verunklart eher die zumeist unter dem Begriff tatsächlich angebotenen technischen Lösungen. Denn es geht dabei nicht allein um Bewältigung von Quantität, sondern auch und vor allem, und das scheint den Innovationsgehalt auszumachen, um die technisch aufwändige Auswertbarkeit ganz unterschiedlich strukturierter Datenbestände für die gebietsübergreifende Datenanalyse. Zeit spielt beim Versprechen Big Data eine kritische Rolle. Schnelligkeit rechtfertigt den Umstieg auf die neuen Lösungen. Zahlreiche Anwendungsbeispiele betreffen deshalb Echtzeit-Analysen von laufenden Datenströmen.

Ansonsten knüpft Big Data aber mehr oder weniger nahtlos an bereits bestehende Konzepte des Data Warehouse und des sog. Data Mining an, die allerdings die vorgehende Zusammenführung und Aufbereitung strukturierter Datenbestände voraussetzen. Ziel auch von Big Data-Datenanalysen ist es, wie

beim Data Mining auf der Grundlage von Abgleichen zwischen zu ganz unterschiedlichen Zwecken erhobenen Datenbeständen statistisch relevante Korrelationen und damit verwertbare Strukturinformationen zu erhalten. Im Kontext von Unternehmen etwa bedeutet das schlicht: Handlungsempfehlungen und Vorhersagen auf der Grundlage von Zahlen, am besten in Echtzeit, zum Beispiel: wie gefällt den Kunden das neu auf den Markt geworfene Produkt X? Muss interveniert werden, um den Erfolg zu sichern, wenn ja, wie? Wenn dabei zusätzlich Daten sozialer Netzwerke angezapft werden, bezeichnet man das als die Einbeziehung „von Außenansichten“ in Unternehmensentscheidungen[6], ganz so, als ob diese Unternehmen vormals keinerlei „Realitätskontakt“ gehabt hätten. Ein weniger kontroverses, weil nicht personenbezogenes Beispiel mag die Datenanalyse des weltgrößten Windradherstellers Vestas sein, der auf der Grundlage unterschiedlichster, aufwändiger Zusammenführung von Informationsquellen die möglichen Erträge eines Windradstandortes prognostiziert.[7]

Drei Vorfälle des zurückliegenden Jahres hatten hinreichend deutlich werden lassen, welche Veränderungen sich mit dem Namen Big Data verbinden können, aber auch welches Konfliktpotential in dieser Entwicklung steckt. Zum einen war dies der Miniskandal um die angedachte Auswertung sozialer Netzwerke durch die SCHUFA. Die größte bundesdeutsche Auskunftsteilnehmerin hatte Informatiker von der Uni Potsdam damit beauftragt zu prüfen, welches Potential in der Auswertung der Daten sozialer Netzwerke für die Bonitätsprüfung stecken könnte. Die öffentliche Kritik an dieser Forschung fiel bei Bekanntwerden jedoch so heftig aus, dass die SCHUFA innerhalb kürzester Zeit die Reißleine ziehen musste und das Projekt beendete.[8] Ähnlich erging es dem Telekommunikationsanbieter Telefonica, der angekündigt hatte, mittels Big Data den finanziellen Nutzen der eigenen Datenbestände, offenbar also vor allem der dem Fernmeldegeheimnis unterfallenden Standort- und Verkehrsdaten ausloten zu wollen. Als selbst Teile der Bundesregierung dagegen Bedenken anmeldete, ruderte man (zunächst) zurück.[9] Und schließlich musste selbst der notorisch unbelehrbare Zuckerberg sein Facebook für den europäischen Markt zumindest modifizieren: die zuvor bereits allgemein aktivierte Gesichtserkennungssoftware zur Erfassung hochgeladener Bilder wurde auf Druck u. a. der EU-Kommission nach Angaben des Unternehmens – vorerst – standardmäßig wieder ausgeschaltet.[10]

Als typische Beispiele für Anwendungsmöglichkeiten von Big Data werden u. a. Verkehrsdaten der Telekommunikation, aber auch Logdateidaten von Webseiten, RFIDProtokolldaten, Verbrauchsdaten im Energiesektor sowie Überweisungsdaten von Banken genannt. Gerade diese im Rahmen von bislang mehr oder weniger eindeutig zweck-gebundenen Prozessen anfallenden Daten sollen nun für weitere Zwecke verfügbar gemacht und kommerziell verwertet werden. Dabei weisen gerade diese Daten mindestens mittelbar auf das konkrete Verhalten eindeutig bestimmbarer Personen, Haushalte oder Personengruppen zurück, sie zählen damit regelmäßig zu den personenbeziehenden Daten im Sinne der Datenschutzgesetze und unterfallen damit dem Anwendungsbereich dieser Gesetze. Die bekannten Beispielfälle deuten oftmals nur an, was tatsächlich möglich wird, und Anwender wollen teilweise anonym bleiben.[11]

Forschung nutzt Big Data in Einzelfällen ebenfalls bereits: Ein Team der Harvard-Universität hat offenbar in Zusammenarbeit mit kenianischen Mobilfunkprovidern Bewegungsprofile von Millionen Kenianern für einen bestimmten Zeitraum erstellt. Die Ergebnisse wurden mit Karten über die Ausbreitung der Malaria abgeglichen mit dem Ergebnis, dass nun menschliche Anteile bei der Ausbreitung als nachgewiesen gelten. Gerade das letztgenannte Beispiel macht klar: selbstverständlich sind außerordentlich wertvolle Anwendungen denkbar, die unter dem Namen Big Data in Zukunft erfolgen könnten. Gleichzeitig wird deutlich, dass auch hier angesichts der mit der individuellen Profilbildung der Bürger einhergehenden Risiken massive ethische wie rechtliche Fragen aufgeworfen werden.

Was ist vom Erkenntniswert von Big Data zu halten?

Selbst wenn mit dem Hype um Big Data schon das Ende der Wissenschaft beschworen wird[12]; die durch

statistische Verfahren ermittelbaren Korrelationen treffen grundsätzlich andere Aussagen über die Wirklichkeit als kausal angelegte Argumentationen. Die Reichweite ihrer Verwertbarkeit scheint damit begrenzt. Und die Ermöglichung der Auswertbarkeit von zu unterschiedlichen Zwecken erfassten und auf unterschiedliche Weise strukturierten bis unstrukturierten Daten gestaltet sich schon IT-technisch äußerst komplex und aufwändig. Der entscheidende Schritt der Auswertung selbst schließlich muss auf herkömmliche Art und Weise erarbeitet werden, insbesondere muss die Fragerichtung und die anschließende Einordnung und Bewertung der Ergebnisse innerhalb vorhandener Prozesse geklärt werden. Es bleibt dabei: Technikeinsatz erfolgt in komplexen sozialen Systemen.[13] Erst deren Ausgestaltung entscheidet insgesamt über den Erfolg der verwendeten Technik. Hier dürften auch weiterhin, verbunden mit finanziellen Fragen, die größten Umsetzungsschwierigkeiten für Anwender liegen, die insgesamt noch verhalten auf den Hype um Big Data reagieren.[14]

Sollen mittels Big Data über die bloß zusammenführende Datenverarbeitung hinaus auch noch sozial verwertbare, womöglich gar im wissenschaftlichen Sinn reliable und valide Informationen abgeleitet werden, bedarf es erheblicher zusätzlicher Anstrengungen, die nicht mittels IT-Technik allein zu bewältigen sind. Auch handelt es sich im Sinne der Statistik oftmals nicht um verwertbare Daten, weil es ihnen an Genauigkeit mangelt.[15] Insbesondere geht entsprechenden wissenschaftlichen quantitativen oder qualitativen empirischen Untersuchungen die Theoriebildung voraus. Problematisch erscheint dabei, dass die weit überwiegende Anzahl der verwendeten Daten von vornherein aus Prozessen stammen, die für ganz andere Zwecke aufgesetzt wurden und damit bereits auf dieser Ebene eine spezifische Selektivität der „Realitätserfassung“ vorliegt. Diese wird verstärkt durch die bei Big Data explizit so gewollte rein technische Zusammenfüllung gänzlich heterogener Datenbestände und die damit verbundene Begrenzung der Erhebungsmethoden. Diese Ausgangslage setzt der Verwertbarkeit von Big Data im Hinblick auf sozialwissenschaftliche Fragestellungen deutliche Grenzen. Begrenzt ist damit allerdings auch das Wohlfahrtsargument der Big Data-Befürworter: hinreichend komplexe Aussagen über „Wirklichkeit“ wie etwa komplexe gesellschaftliche Abläufe setzen einfach mehr voraus als schiere Quantität und werden auch in Zukunft aufwändiger gestaltet sein. Damit muss nicht ausgeschlossen werden, dass wir möglicherweise dank Big Data mit einer zunehmenden Kultur der (bloßen) Korrelationsargumentation zu rechnen haben, die sich über bestehende Erwartungen hinsichtlich kausaler Argumentation und damit auch über herkömmliche wissenschaftliche Begründungsmuster hinwegsetzen wird. Allerdings wird sich diese stets dem Einwand ihrer begrenzten Aussagekraft ausgesetzt sehen.

Deutlich wird hingegen, dass Big Data-Konzepte Ausdruck einer Methodik sind, wie sie in der Wirtschaftsinformatik bereits länger Anwendung findet und in erster Linie auf die praktische Anwendung in betriebswirtschaftlichen Zusammenhängen abzielt. Die begrenzt verwertbaren Aussagen können, etwa in Form von Tendenzaussagen und statistisch unterlegten Vermutungen in bestimmten Entscheidungsprozessen möglicherweise durchaus Mehrwerte bieten. Sie zielen mehr auf die effiziente Verwaltung von Risiken angesichts einer Fülle von Daten als auf Erkenntnis an sich.

Big Data und Datenschutz

Big Data deutet darauf hin, dass das mittlerweile erreichte Ausmaß auch der personenbeziehbaren Datenerfassung mit dem Ideal einer der Gründerväter des Datenschutzes Spiros Simitis, der so gepriesenen „Datenaskese“, nicht mehr zusammen zu bringen ist. Das blanke Gegenteil ist der Fall. Ein Gros der heute oftmals erhobenen Daten fällt zu-dem automatisiert an und erfüllt bei den verantwortlichen Stellen, außer der primären Nutzung, ansonsten keinerlei Funktion. Aus dieser datenschutzrechtlich an sich eindeutig auf die Pflicht der Löschung verweisenden Situation macht die IT-Industrie eine Tugend und verspricht das Heben eines verborgenen Schatzes von Antworten aus dem Daten-Durcheinander. Genau an der Idee von Big Data wird das angesichts mehrerer potcom-Blasen etwas aufgesetzt begeistert wirkende Marketinggeklingel von Daten als dem „Öl des 21. Jahrhunderts“ nachvollziehbar.

Das Versprechen Big Data geht allerdings mit zentralen Prinzipien des Datenschutzes nicht konform. Denn es suggeriert zukünftig Datenverarbeitern, getreu dem Motto „wer weiß, wozu es gut sein wird“, dass es im Hinblick auf etwaige spätere Auswertungsmöglichkeiten geradezu ökonomisch falsch wäre, die einmal erfassten Daten wieder zu löschen, zu unterschiedlichen Zwecken getrennt zu verarbeiten oder, noch davorliegend, überhaupt begrenzende Festlegungen hinsichtlich der zu erhebenden Daten selbst vorzunehmen. Selbstverständlich hat es solche Überlegungen auch in anderen Zusammenhängen immer schon gegeben. Mit dem Konzept von Big Data aber verschärft sich dieser Interessengegensatz deutlich. Wenn im Kern zahlreiche Anwendungen von Big Data die noch umfassendere Bildung von computergenerierten Verhaltens-, Persönlichkeits- oder Bewegungsprofilen zum Ziel haben, verstärkt dies den Druck auf die davon Betroffenen. Mehr Entscheidungsprozesse werden aufgrund der Bewertung der Zugehörigkeit zu einer statistisch relevanten Vergleichsgruppe und den für diese Gruppe festgelegten Annahmen getroffen werden. Bei dieser auch und gerade gruppenspezifischen Diskriminierung drängen sich aus rechtlicher Sicht Fragen möglicher ungerechtfertigter Ungleichbehandlung auf. Hier wird erkennbar, dass über die Zielsetzung des Datenschutzes hinaus weitere wichtige auf Ausgleich abzielende Gemeinwohlinteressen berührt sein können.[16]

Eines der Ziele des Datenschutzes besteht darin, Technikeinsatz menschengerecht zu gestalten. Das bedeutet im demokratischen Rechtsstaat Transparenz und Gestaltbarkeit der Technik sowie die Gewährleistung der Rechte der Betroffenen. Soweit Big Data keine personenbezogenen oder personenbeziehbaren Datenbestände umfasst, mögen Datenschutzfragen im herkömmlichen Sinne keine Rolle spielen. Gleichwohl werden auch dann Fragen der Informationsordnung berührt, wie sie bereits angedeutet wurden: in welchen Institutionen und Entscheidungsverfahren erscheint es gesellschaftlich wünschenswert, die geschilderte spezifische Selektivität von Big Data-Verfahren anzuwenden?

Doch in den allermeisten gegenwärtig diskutierten Anwendungsbeispielen geht es gerade um die Auswertung zumindest personenbeziehbarer Datenbestände. Denn stets drehen sich diese Anwendungen um das Verhalten von Menschen. Mal sollen Kaufvorlieben vorhergesagt, mal das Risiko der fehlenden Rückzahlung eines Kredites (kreditorisches Ausfallrisiko), mal das Risiko des kriminellen Verhaltens evaluiert werden, oder es soll die Möglichkeit des Aufspürens eines Straftäters sichergestellt werden. Dass für Big-Data-Auswertungen dabei womöglich Klarnamen oder auch andere identifizierende Merkmale weggelassen werden (Pseudonymisierungen) und das Erkenntnisinteresse der Datenverarbeiter von vornherein nicht auf einzelne Personen, sondern auf aggregierte Datenbestände und allgemeinere Aussagen abzielt, ändert an der grundsätzlichen Anwendbarkeit der Datenschutzgesetze nichts. Denn zum einen sind auch weitgehend pseudonymisierte Datenbestände in aller Regel auf einfache Weise re-personalisierbar. Die entsprechenden grundlegenden Untersuchungen dazu erfolgten bereits vor Jahrzehnten. Folglich hätte der Hack eines unzureichend gesicherten, bloß pseudonymisiert vorgehaltenen Big-Data-Datenbestandes genau so schwere Folgen für die betroffenen Einzelnen wie sonst auch, nur mit einer erheblich größeren Anzahl von Betroffenen. Zum anderen haben auch die zunächst auf größere Gruppen und statistisch relevante Größen abzielenden Auswertungen spätestens dann ganz konkrete Auswirkungen auf einzelne Betroffene, wenn diese den ermittelten Risikogruppen zugeordnet und den damit angeordneten Folgen unterliegen oder eben auch nicht (Unterbreitung einer Werbung oder eines Vertragsangebots; Angebot vergünstigter Konditionen; Kündigung des Vertrages; Nichtaufnahme einer Vertragsbeziehung usw.). Im staatlichen Bereich stellt sich dies grundsätzlich mit anderen, oftmals gravierenderen Handlungsfolgen dar, wenn auch nicht pauschal mit einer höher zu bewertenden Grundrechtsintensität.

Der Datenschutz stellt in seiner bisherigen gesetzlichen Konzeption deutlich auf den Schutz des Individuums als Grundrechtsträger ab. Ziel ist es danach, mögliche entstehende Nachteile durch die Verarbeitung der die Bürgerinnen und Bürger betreffenden Daten und Informationen möglichst frühzeitig und umfassend zu verhindern. Big Data mit der je nach Anwendungsfeld eigenen Mischung aus verstärkter, wenn auch nicht unbedingt physikalischer, sondern der vernetzten Zusammenführung von unterschiedlichen Datenbeständen und deren Auswertbarkeit verstärkt den aufgrund technischer wie gesellschaftlicher Entwicklungen zentralen Prinzipien bereits bestehenden Druck auf zentrale Prinzipien des Datenschutzes. Dies gilt sowohl für den öffentlichen als auch nichtöffentlichen Bereich. Allerdings dürfte der öffentliche Sektor sowohl aufgrund langfristig finanziell klammer Lage (auch im Sicherheitsbereich) als auch einer recht ausdifferenzierten

Verfassungsrechtsprechung weniger im Vordergrund stehen als die Privatwirtschaft.[17]

Wie bereits erwähnt, wird es für Datenverarbeiter mit Blick auf die Versprechen der Big-Data-Anbieter noch weniger verständlich werden, weshalb überhaupt und vorab eine möglichst konkrete Festlegung hinsichtlich des Zweckes einer Datenverarbeitung erfolgen sollte. Schließlich kann ja erst am Ende gesagt werden, ob nicht doch wichtiger Zusatznutzen aus den Datenbeständen zu ziehen war. Der mit dem Grundsatz der Zweckfestlegung eng verbundene Grundsatz der Erforderlichkeit, mit dem eine weitere Strukturierung und Begrenzung der Verwendungszusammenhänge personenbezogener Daten mit Blick auf die jeweilige Sachaufgabe erfolgen soll, kann dann nur eine geringere steuernde Wirkung entfalten. Auch die mit Zweckfestlegung und Erforderlichkeit zusammen begründeten Löschpflichten für bestehende Daten dürften in vielen Fällen unter Verweis auf die mit Big-Data Instrumenten möglichen Auswertungen auch scheinbar nicht erforderlicher Datenbestände eher schwieriger durchsetzbar werden.

Mögliche absolute Grenzen von Big Data-Auswertungen können zum einen in der Art der ausgewerteten Datenbestände selbst als auch in der Verdichtung der entstehenden Persönlichkeits- und Verhaltensprofile liegen. Die Empörung über die geplante Telefonica-Datenauswertung weist insoweit in die Richtung des ersten Falles, als zumindest Verkehrsdaten wie auch Inhaltsdaten unstreitig dem Fernmeldegeheimnis und damit auch einer engeren Zweckbestimmung unterliegen als andere Daten. Die mit Big-Data-Verarbeitungen zumeist verbundene, rechtfertigungsbedürftige Zweckänderung kann möglicherweise nur aufgrund konkreter gesetzlicher Ermächtigungen geöffnet werden. Ein Beispiel für die kritische Verdichtung von Profilen können Auswertungen von Sozialen Netzwerken bilden, selbst wenn diese als allgemein „öffentlich zugänglich“ wahrgenommen werden. Die unter Berufung auf die Öffentlichkeit von Datenbeständen beanspruchte Privilegierung bei der rechtlichen Bewertung der Datenverarbeitung verkennt insoweit die entstandene Vielfalt ganz unterschiedlicher Öffentlichkeiten und die jeweils damit verbundenen, generalisierbaren Erwartungen der Betroffenen. Hier muss der Gesetzgeber konkret nachbessern.

Eine erneute Betrachtung verdient angesichts von Big Data der Grundsatz der Richtigkeit der Daten. In dem Umfang, wie Profile von Verbrauchern/Bürgern mit vage bleibenden Korrelationsfaktoren angereichert werden, die zudem angesichts zunehmend selbstlernender Algorithmen kaum mehr näher erläutert werden können, entstehen hochselektive und nicht mehr rekonstruierbare Bilder von Personen.

Im Vordergrund der datenschutzrechtlichen Bewertung muss wie gesagt auch stehen, dass mit Big Data eine Verstärkung der im Zusammenhang mit der Data Warehouse/Data Mining-Debatte bereits bekannten Zuordnung von Einzelnen zu statistisch ermittelten Gruppen erfolgt und damit in den unterschiedlichsten Entscheidungsprozessen einseitig eine Bewertung der Betroffenen über ihre Gruppenzugehörigkeit erfolgt. Wenn über diese Zuschreibungen benachteiligende Festlegungen erfolgen, sind weit über den Einzelfall hinaus Gemeinwohlinteressen berührt. Es fragt sich, ob die bislang bestehenden gesetzlichen Vorgaben speziell für das sog. Kreditscoring mit der Festlegung der manuellen Letztbearbeitung und -entscheidung durch tatsächliche Personen ausreichen, um hier weitreichende automatisierte Entscheidungsprozesse zum Nachteil der Bürger zu verhindern.

Diese wie auch die weiteren offenen Datenschutzfragen lassen sich nur dann überzeugend lösen, wenn Big-Data-Anwendungen für personenbeziehbare Datenbestände mit verlässlichen, d. h. grundsätzlich nachvollziehbaren und unabhängig geprüften Anonymisierungsverfahren durchgeführt werden.[18]

Big Data und die EU-Datenschutzreform

Die Antwort des Datenschutzes auf diese Entwicklungen entscheidet sich derzeit vorrangig auf europäischer Ebene in der Debatte um die EU-Datenschutzreform. Der ambitionierte Vorschlag der EU Kommission erfährt zurzeit eine umfassende Ergänzung und Änderung durch das EU-Parlament. Bei aller Vorläufigkeit dieser voraussichtlich noch einige Zeit andauernden Reformdebatte deuten komplexe Big Data-

Anwendungen an, worauf es ankommen sollte. Zum einen kann angesichts der künftigen Ubiquität auch solcher Techniken wie Big Data Datenschutz nur mit präventiver Zielrichtung effektiv bleiben. Es bleibt deshalb richtig, auf eine Vorverlagerung des Schutzes in die Technik-ebene selbst zu dringen. Was das konkret für Big Data heißen mag bleibt zu erforschen.[19] Jedenfalls könnte es bereits auf Herstellerebene um die Transparentmachung von Big-Data-Zugriffen auf heterogene Datenbestände gehen, aber auch konkrete Auskunftsverfahren für Betroffene könnten technisch unterstützt werden. Die gesetzgeberische Herausforderung besteht darin, dieses sog. Privacy by Design wenn nicht durch direkte gesetzliche Verpflichtungen, dann zumindest mittelbar über privatrechtliche Haftungsnormen durchzusetzen. In der Vielfalt der Anbieter wird es für einsetzende Unternehmen wie auch Verbraucher von Interesse sein, ob es sich um vertrauenswürdige Unternehmen und Instrumente handelt, so dass auch weiche Steuerungsinstrumente des Datenschutzes wie Gütesiegel und Auditierungen eine deutliche Stärkung in der EU-Reform erfahren sollten.

Alle diese Überlegungen können aber nicht darüber hinwegtäuschen, dass es insbesondere bei der Regelung der Profilbildung, der Einwilligung und den Transparenzbestimmungen deutlicher Verbesserungen zugunsten der Bürger bedarf. Die bestehenden Entwürfe der Reform genügen noch nicht, vor allem sind sie in vielen Punkten zu unkonkret. Absolute Grenzen der Zulässigkeit müssen im Sinne etwa des Diskriminierungsschutzes festgelegt werden, damit nicht über den Weg abgenötigter Einwilligungen einzelner Betroffener Gemeinwohlziele ausgehebelt werden können. Einwilligungen selbst müssen entsprechend aufwändig geregelt werden, um Freiwilligkeit sicherzustellen, echte Alternativen offen zu halten und informierte Entscheidungen zu gewährleisten. Verhindert werden muss angesichts der informatorischen wie kognitiven Überforderung der Verbraucher insbesondere, dass über einmal erteilte Pauschaleinwilligungen die bleibende Zusammenführung ganzer Konzerndatenbestände zur gängigen Praxis wird. Soweit Big Data die Zusammenführung umfänglicher Datenbestände ermöglicht, sind entsprechende Verstärkungen der im Rahmen des Datenschutzes zu gewährenden Datensicherheit grundlegend. Auf die verstärkte Bedeutung der Datensicherheit haben insoweit auch die Entscheidungen des Bundesverfassungsgerichts zur Vorratsdatenspeicherung sowie zum Grundrecht auf Integrität und zur Vertraulichkeit informationstechnischer Systeme hingewiesen. Die bislang vorgesehenen Bestimmungen des Entwurfs einer EU-Datenschutzgrundverordnung bieten hierzu wenig Sachgerechtes. Die möglichen Antworten auf die u. a. mit Big Data verbundenen Herausforderungen zeigen, dass in diesem Bereich die Hoffnung auf weniger Regulierung weder realistisch noch sachgerecht erscheint, wenn man das Ziel einer möglichst gemeinwohlverträglichen- und verfassungskonformen Gestaltung des Einsatzes von IT-Technologie auch angesichts von Big Data nicht aufzugeben bereit ist.

[1] Begriff nach A. Grunewald, TAB-Brief Nr. 39, S. 6, 2011, abrufbar unter http://www.tab-beimbundestag.de/de/pdf/publikationen/tab-brief/TAB-Brief_039.pdf

[2] Ein typisches Beispiel im Kontext von Big Data bietet der Aufsatz von C. Anderson, der das „Ende der Wissenschaft“ durch Big Data beschwört:

http://www.wired.com/science/discoveries/magazinellfr07/pb_theory

[3] Vgl. TAB-Brief, a.a.O.

[4] Hier sind es die Internetnutzerinnen selbst, die die Daten hinterlassen, worauf J. Hoffmann hinweist, vgl. Magazin der Böll-Stiftung, Ausgabe 2/ 2012, S. 33.

[5] Definition des Arbeitskreises Big Data, vgl. <http://www.bitkom.org/de/wirueberuns/70822.aspx> (21.12.2012).

[6] Vgl. etwa den Beitrag von V. Markl, <http://www.t-systems.de/news-media/gastbeitrag-voninformatiker-volker-markl-bigdata-und-der-schmale-grat-zwischen-rechtzeitig-und-zu-spaet/> 995378

[7] Die Zeit, 13.01.2013, Schwerpunkt Big Data.

[8] Süddeutsche vom 8. Juni 2012, <http://www.sueddeutsche.de/digital/kritik-an-neuem-scoring-verfahren-facebook-projekt-der-schufa-abgeblasen-1.1377327>

[9] <http://www.tagesschau.de/Wirtschaft/telefonica-deutschland104.html>

[10] <http://www.sueddeutsche.de/digital/erfolg-fuer-datenschuetzer-facebook-schaltet-gesichtserkennung-ab-1.1474966>

[11] Vgl. z.B. die Studie der BITKOM, abrufbar unter <https://www.bitkom.org/de/publikationen13833773446.aspx>

[12] Vgl. Fn. 2.

[13] Gutes Beispiel bei W. Rammert, S. 342 in: Bild-Raum-Kontrolle, hrsg. durch L. Hempel und J. Metelmann, 2005.

[14] Auch wenn die Verkaufszahlen von Big Data-Produkten auffällige Zuwächse verzeichnen, wird damit noch keine Aussage über den tatsächlichen Nutzen dieser Technologie einschließlich der Frage nach ihren Nebenwirkungen getroffen.

[15] So auch zum Data Mining allgemein: <http://de.wikipedia.org/wiki/Data-Mining>

[16] In diese Richtung auch Schallaböck, Sonderheft Digitale Demokratie der Böll-Stiftung, a.a.O., S. 34, der durch Big Data Wissensmonopole Privater entstehen sieht.

[17] Wenn auch die genannten EU-geförderten Projekte wie INDECT oder auch die von der EU-Kommission vorangetriebene Rasterfahndung in EU-Flugpassagierdaten oder das Beharren auf der Vorratsdatenspeicherung von TK-Verkehrsdaten durchaus Anlass zur Sorge geben, vgl. dazu den jährlich erscheinenden Grundrechte-Report der Bürgerrechtsorganisationen.

[18] Dass dafür die Ressourcen in Unternehmen vorhanden sein sollten, darauf weist zutreffend J. P. Albrecht im Interview in „Schwerpunkt: Big Data“ hin, Die Zeit vom 13.01.2013.

[19] Die Datenschutzforschung zeichnet sich als ein weiter wachsendes Feld innerhalb der Sicherheitsforschung ab, vgl. u.a. die Fraunhofer-Gesellschaft, die sich mit den wichtigen Anonymisierungsverfahren befasst, Die Zeit vom 13.01.2013, Schwerpunkt: Big Data.

<https://www.humanistische-union.de/thema/big-data-eine-neue-herausforderung-fuer-datenschutz/>

Abgerufen am: 24.04.2024