Informationsfreiheit und Informationsrecht

Jahrbuch 2015

Herausgegeben von
Alexander Dix
Gregor Franßen
Michael Kloepfer
Peter Schaar
Friedrich Schoch
Andrea Voßhoff
und der Deutschen Gesellschaft für Informationsfreiheit

Der juristische Verlag
lexxion
Berlin
Brüssel

Informationsfreiheit und Informationsrecht

Jahrbuch 2015

Herausgegeben von
Alexander Dix
Gregor Franßen
Michael Kloepfer
Peter Schaar
Friedrich Schoch
Andrea Voßhoff
und der Deutschen Gesellschaft für Informationsfreiheit

Der juristische Verlag
lexxion
BERLIN
BRÜSSEL

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Das Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, der Entnahme von Abbildungen, der Funksendung, der Wiedergabe auf fotomechanischem oder ähnlichem Wege und der Speicherung in Datenverarbeitungsanlagen, bleiben vorbehalten.

Das Werk wurde mit größter Sorgfalt zusammengestellt, dennoch übernimmt der Verlag keine Haftung für inhaltliche und drucktechnisch bedingte Fehler.

Zitierhinweis: JB InfoR

ISBN Print: 978-3-869 65-281-8
ISBN E-Book: 978-3-869 65-282-5

© 2016 Lexxion Verlagsgesellschaft mbH · Berlin
www.lexxion.de

Satz: Susanne Bonnet, Berlin
Umschlag: TOZMAN Satz & Grafik, Berlin
Foto: fotolia.com © Weissblick

Vorwort

Die Informationsfreiheit und das Informationsfreiheitsrecht waren auch 2015 in Bewegung. Dies zeigt nicht nur die umfangreiche Rechtsprechung. Auch die Rechtswissenschaft ist weiterhin damit beschäftigt, Anwendungsbereich und Grenzen, aber auch Optimierungsmöglichkeiten der Informationsfreiheit zu klären. Transparenz ist gerade auch da geboten, wo Entscheidungen zunehmend an Maschinen delegiert werden, die Big Data unter Einsatz von Algorithmen auswerten. Und in der „Nachbarschaft" des Informationsfreiheitsrechtes sind grundlegende Neuerungen erkennbar: Das europäische Datenschutzrecht durchlebt mit der jüngst verabschiedeten Datenschutz-Grundverordnung einen Generationenwechsel. Der EuGH hat 2015 mit den Entscheidungen zu Safe Harbor und Google-Spain wichtige Weichen gestellt.

All dies und manches mehr soll das vorliegende Jahrbuch widerspiegeln:

In Baden-Württemberg wurde kurz vor Redaktionsschluss ein Informationsfreiheitsgesetz beschlossen. Rheinland-Pfalz vollzieht den Wechsel in Richtung eines Transparenzgesetzes. Über die Entwicklung in beiden Ländern und über erste Überlegungen in Niedersachsen und Hessen informiert der Beitrag von **Stefan Brink** und **Sonja Wirtz.**

Christoph Gusy verortet die Informationsfreiheit im verfassungstheoretischen und verfassungsrechtlichen Kontext, greift dabei die Grundfrage der Transparenz als Eigen- oder (nur) instrumenteller Wert auf und sieht die im Demokratieprinzip wurzelnde Transparenz „als ein im Postulat effizienter Verwaltungstätigkeit mitgedachtes und angelegtes Sekundärphänomen". Transparenz versteht Gusy als Ergebnis von Abwägungen und Aushandlungsprozessen. *„In diesem Sinne ist auch der transparente Staat kein gläserner Staat.* Er muss es nicht sein und er darf es auch nicht sein. Eine derartige Forderung kann allenfalls ein politisches Anliegen sein. Eine verfassungstheoretische oder -rechtliche Begründung findet sich nicht – ebenso wenig wie für den alten ‚geheimen Staat'."

Terry Martin, als amerikanischer Journalist seit vielen Jahren in Berlin tätig, nimmt den „Freedom of Information Act" und das vergleichsweise junge IFG mit ihren durchaus unterschiedlichen historischen Wurzeln in den Blick.

Peter Schaar greift mit seinem Beitrag zur Algorithmentransparenz ein zunehmend bedeutendes Phänomen auf, das einer umfassenden Untersuchung bedarf. Automatisierte Entscheidungen nehmen deutlich bis dramatisch zu. Sie stützen sich häufig auf korrelierende Variablen oder Verfahren zur Muster-

erkennung in großen Datenbeständen. Der Algorithmus beschreibt die zugrundeliegende Methodik. Transparenz ist hier geboten, de lege lata aber nicht gewährleistet.

Moritz Lebsanft analysiert die – teils inhomogene – Rechtsprechung des EuGH bzw. des EuG und die nationale Judikatur zu Fällen, in denen gemeinschaftsrechtliche oder im deutschen Recht geregelte Informationszugangsrechte auf der einen und ebenfalls gemeinschaftsrechtlich vorgegebene Berufsgeheimnisse oder der Informantenschutz auf der anderen Seite in Ausgleich zu bringen sind.

Die Risiken transatlantischer Kommunikation sind unter anderem durch die Veröffentlichungen von Edward Snowden in den Fokus des öffentlichen Interesses geraten. Der massive Vertrauensverlust bei Bürgerinnen und Bürgern, aber auch bei Unternehmen belastet die Verhandlungen der EU mit den USA (u.a.) über die Transatlantic Trade and Investment Partnership (TTIP) und die multilateralen Verhandlungen über das Trade in Services Agreement (TISA). In seinem Beitrag macht **Alexander Dix** deutlich, „dass der Mangel an Transparenz der TTIP-Verhandlungen zur Achillesferse für die transatlantischen Freihandelsgespräche zu werden droht." Dix plädiert dafür, hohe europäische Datenschutzstandards auch weiterhin zu verteidigen und sie – wie im GATS-Abkommen aus dem Jahre 1994 ausdrücklich festgelegt – auch weiterhin nicht als abzubauende nicht-tarifäre Handelshemmnisse anzusehen.

Sisyphus müssen wir uns als einen glücklichen Menschen vorstellen: **Henning Blatt** hat auf fast 80 Seiten eine Fülle von Gerichtsentscheidungen zum Informationsrecht von Bund und Ländern ausgewertet und damit auch diesmal einen unverzichtbaren „Klassiker" beigesteuert.

Der presserechtliche Auskunftsanspruch war in den letzten Jahren mehrfach Gegenstand mitunter überraschender höchstrichterlicher Entscheidungen. Auch wenn die landespresserechtlichen Auskunftsansprüche eine große Homogenität aufweisen, sieht **Holger Greve** hier verfassungsrechtlich eröffneten Spielraum für eine Stärkung des Informationszuganges durch Reduzierung von Beschränkungen.

Hansjürgen Garstka beleuchtet die Stellung der bzw. des Bundesbeauftragten für den Datenschutz und die Informationsfreiheit nach der – durch die Rechtsprechung des EuGH gebotenen – Novellierung des BDSG. Nach seiner Auffassung sind damit „zwar die wesentlichen Mängel im Hinblick auf die ‚völlige' Unabhängigkeit beseitigt". Unter anderem mit Blick auf die Einschränkung der Berechtigung der Bundesbeauftragten zur Zeugenaussage und das (weiterhin nicht explizit gewährte) Recht, aus eigener Initiative vor dem Parlament oder Ausschüssen zu erscheinen und zu reden, sieht er noch Optimierungsmöglichkeiten. Mit der gesetzlichen Aufgabenzuweisung in § 12 Abs. 2 setzt das IFG die – mit der

Novellierung des BDSG weiter gestärkte – Unabhängigkeit der Bundesbeauftragten voraus. Mit der Novellierung des BDSG sehe ich deshalb auch die Position der bzw. des Bundesbeauftragten für die Informationsfreiheit gestärkt.

Gemeinsam mit **Sven Hermerschmidt** gebe ich einen Überblick über die Datenschutz-Grundverordnung, auf deren endgültigen Inhalt sich die europäischen Institutionen jüngst geeinigt haben. Der Beitrag befasst sich mit dem Verhältnis zwischen dem Grundrecht auf Datenschutz und dem freien Zugang zu Informationen, nicht nur bezogen auf das Informationsfreiheitsrecht, sondern vor allem im Hinblick auf das Recht der freien Meinungsäußerung. Ein weiterer Schwerpunkt liegt auf einer Darstellung der Transparenzvorschriften zugunsten der datenschutzrechtlich Betroffenen.

Matthias Bergt kommentiert die Entscheidungen des EuGH zu Safe-Harbor und Google-Spain. Aus seiner Sicht „scheint sich der EuGH als Bundesverfassungsgericht 2.0 zu etablieren, das die Grundrechte besser schützt als das Original".

Der Beitrag von **Dennis-Kenji Kipker** beschäftigt sich mit den verfassungsrechtlichen Anforderungen an den Einsatz sicherheitsbehördlicher Verbunddateien und dabei auch mit der Verbesserung von Transparenz, Dokumentation und Kontrolle. Abgerundet wird dieses Jahrbuch durch die Überlegungen von **Ludwig Maidowski** zur Entscheidung des EuGH zur privaten Videographie im Grenzbereich des öffentlichen und des privaten Raumes und den Beitrag von **Hannah Wirtz** zur Umsetzung der konsolidierten PSI-Richtlinie und zur Weiterverwendung kultureller Informationen.

Auch dieses Jahrbuch spiegelt damit die vielfältigen Facetten und Bezüge der Informationsfreiheit und ihres rechtlichen Umfeldes wider.

Andrea Voßhoff

Inhalt

*Peter Schaar**

Algorithmentransparenz

Inhaltsübersicht

I. Einleitung

Big Data-Ansätze, die auf der Sammlung, Zusammenführung und Auswertung großer Datenmengen beruhen, beinhalten erhebliche Risiken für das Recht auf informationelle Selbstbestimmung. Sie sind mit den bisherigen, auf die jeweiligen Daten fokussierten datenschutzrechtlichen Auskunftsansprüchen nicht hinreichend zu beherrschen. Rechtliche Verpflichtungen zur Offenlegung der bei der Verarbeitung und Bewertung verwendeten Algorithmen können dazu beitragen, diskriminierende Praktiken zu erkennen und negative Auswirkungen des Profilings zu vermeiden.

II. Die Big Data-Gesellschaft

Nach den von Gordon Moore 1965 formulierten Erkenntnissen verdoppelt sich die Verarbeitungskapazität elektronischer Komponenten regelmäßig („Moore'sches Gesetz") bei unveränderten Komponentenkosten.[1] So verdoppeln sich die Kapazitäten von Speichermedien, die Verarbeitungsgeschwindigkeit von Prozessoren alle 18 bis 24 Monate – bei unverändertem Preis. Die IT-Akzeleration hat dramatische Folgen: Bereits heutzutage haben elektronische Komponenten analoge Systeme in nahezu allen Bereichen der Kommunikations-, Antriebs-, Mess- und Steuerungstechnik ersetzt. Informationen werden heute fast durchgängig digital erfasst, gespeichert, übertragen und ausgewertet. Die digital gespeicherte Informa-

* Peter Schaar ist Vorsitzender der Europäischen Akademie für Informationsfreiheit und Datenschutz und war von 2003 bis 2013 Bundesbeauftragter für Datenschutz und Informationsfreiheit.

1 *Moore*, Cramming More Components Onto Integrated Circuits, Electronics, 1965, S. 114.

tionsmenge übersteigt den Umfang der konventionell gespeicherten Informationen um ein Vielfaches.[2] Rein mengenmäßig spielen analog erfasste Informationen kaum noch eine Rolle.

Zunehmende Bedeutung gewinnen dabei automatisch generierte Daten, die von Sensoren geliefert werden und bei jeder digitalen Transaktion anfallen (Metadaten). Gebrauchsgegenstände, Stromzähler und andere Messgeräte, Verpackungs- und Verbrauchsmaterialien, Tickets für Konzert- und Sportveranstaltungen und andere Dokumente werden mit Computerchips ausgestattet. Über die Luftschnittstelle auslesbare RFID-Chips sind ein wichtiges Verbindungsglied für die Einbindung von Gegenständen in Netzwerke, für die Verbindung von realer und virtueller Welt (Internet of Things – IoT). Auch andere Techniken gestatten den Zugriff auf eingebaute Elektronik, etwa Bluetooth und WLAN-Schnittstellen, die sich mit lokalen Funknetzen verbinden. Die grundlegende Idee des IoT besteht darin, alle möglichen Gegenstände in die Infrastruktur des Internets zu integrieren. Dies ermöglicht nicht nur die elektronische Kontrolle der jeweiligen Geräte – etwa die Fernmessung der Temperatur in einem Kühlhaus oder die Feststellung des Standorts eines Kfz über das Internet. Durch den Rückgriff auf die nahezu unbegrenzten Informationsressourcen des Internets können vergleichsweise einfache elektronische Geräte „smarte" Funktionalitäten erhalten. Der Kühlschrank, der auf einem Display den aktuellen Wetterbericht anzeigt, misst nicht die Außentemperatur und die Luftfeuchtigkeit, sondern bezieht seine Informationen aus dem Netz. Durch tausende von „Apps" (Applications) wird aus dem Handy ein Smartphone, das auf nahezu unbegrenzte Informationsressourcen im Internet zugreifen kann.

Im Jahr 2015 waren bereits mehr als 15 Milliarden Gegenstände per Funk miteinander vernetzt und es ist zu erwarten, dass sich diese Zahl bis 2020 mindestens verdreifachen wird. Aus dem „Internet of Things" wird auf diese Weise ein „Internet of Everything".[3] Die allgegenwärtige Vernetzung geht dabei über eine bloß quantitative Ausweitung der Datenverarbeitung hinaus. Wenn etwa Häuser und Wohnungen „intelligent" werden und sich auf die Gewohnheiten der Bewohner einstellen, ändert sich auch unsere Wahrnehmung und unser Verhalten. Die Grenze zwischen „wirklicher" und „virtueller" Realität verschwindet zusehends. Wissenschaftler sprechen in diesem Zusammenhang von „Connected Spaces", in denen die verschiedensten Geräte, Gegenstände und Personen in ständiger Kommunikation stehen, sich gegenseitig anpassen und voneinander lernen. Der

2 *Dutcher*, Data Size Matters [Infographic], 6.11.2013, abrufbar im Internet unter http://datascience.berkeley.edu/big-data-infographic/ (letzter Zugriff 31.12.2014).

3 *Baker*, Connected spaces: the next step for the internet of things, The Guardian Online, 5.2.2015, abrufbar im Internet unter http://www.theguardian.com/media-network/2015/feb/05/connected-spaces-should-be-the-next-step-for-the-internet-of-things (letzter Zugriff 3.11.2015).

Einbau von IT in Kleidung deutet einen weiteren Sprung an: Sensoren rücken immer näher an unseren Körper heran. Bereits heute werden in manchen Kliniken Neugeborene mit einem RFID-Armband versehen, um so Verwechslungen im Krankenhaus zu verhindern. Ähnliche Projekte gibt es auch für Demenzkranke, die man auf diesem Wege daran hindern möchte, ihr Krankenhaus oder ihre Wohnanlage zu verlassen;[4] eine Art „elektronische Fußfessel" für Alte. Sehr gut verkaufen sich Fitnessarmbänder, die nicht nur die Schritte der Träger zählen, sondern auch deren Schlafgewohnheiten, bisweilen sogar andere Vitalfunktionen und den Aufenthaltsort aufzeichnen und basierend auf der Analyse so gewonnener Daten ihren Trägern Tipps für ein gesünderes Leben liefern. Dass es für die so generierten Angaben auch andere Interessenten gibt, ist naheliegend. So haben Versicherungen angekündigt, ihren Kunden einen Rabatt zu gewähren, wenn sie einen gesunden Lebensstil pflegen. Zum Nachweis sollen Fitness-Tracker und „Gesundheits-Apps" eingesetzt werden.[5] Der nächste logische Schritt sind in den Körper eingebaute Funkchips, die zur Ortung, zur Erfassung medizinischer Werte oder auch zur Identifikation verwendet werden.

„Big Data" steht wie kein anderer Begriff für den Übergang zu einem neuen Modell des Umgangs mit Informationen. Der Begriff umschreibt den Umgang mit riesigen Datenmengen, „die zumeist im Rahmen einer Zweitverwertung zusammengeführt, verfügbar gemacht und ausgewertet werden."[6] Bisweilen wird auch von den „3 V's" gesprochen: „high-volume, high-velocity and high-variety information assets".[7]

Wesentliche Voraussetzungen für Big Data sind bereits vorhanden: Sensoren liefern eine immer größere Informationsbasis, das Internet verknüpft die Datenmassen und immer leistungsfähigere Speicherchips ermöglichen Transaktionen mit zunehmend größeren Datenmengen ohne wahrnehmbare zeitliche Verzögerung. Daten aus unterschiedlichsten Quellen werden zusammengeführt und ggf. in Echtzeit auf Auffälligkeiten untersucht.

---

4 *Gnuess*, Funkchips für das Krankenhaus der Zukunft, Handelsblatt, 10.11.2006, abrufbar im Internet unter http://www.handelsblatt.com/technologie/forschung-medizin/medizin/rfid-funkchips-fuer-das-krankenhaus-der-zukunft-seite-all/2731124-all.html (letzter Zugriff 6.1.2015).

5 Vgl. *Janker*, Wir werden manipulierbar und unfrei – Juli Zeh über das Generali-Modell, Süddeutsche Zeitung, 26.11.2014, abrufbar im Internet unter http://www.sueddeutsche.de/kultur/juli-zeh-ueber-das-generali-modell-wir-werden-manipulierbar-und-unfrei-1.2232147 (letzter Zugriff 25.2.2015).

6 *Weichert*, Big Data – eine Herausforderung für den Datenschutz, in: Geiselberger/Moorstedt (Hrsg.), Big Data – Das neue Versprechen der Allwissenheit, 2013, S. 133.

7 President's Council of Advisors on Science and Technology (PCAST), Big Data and Privacy: A Technological Perspective, May 2014, S. 2.

III. Algorithmische Klassifikation

Bei der automatisierten Informationsverarbeitung werden Daten nach in der Hard- und Software implementierten Regeln verarbeitet, den sogenannten Algorithmen. Die klassischen Small Data-Algorithmen orientieren sich an der jeweiligen Aufgabe. Letztlich geht es um möglichst effiziente Verfahren, mit denen sich aus einer definierten Datenmenge ein Ergebnis erzielen lässt – etwa die Abwicklung der Gehaltszahlung. Welche Daten erforderlich sind, ergibt sich aus der jeweiligen Aufgabe – in unserem Beispiel: Arbeitsstunden, Tarifgruppe, individuelle Zulagen. Niemand würde auf die Idee gekommen, zur Gehaltsberechnung Daten über das Wetter, das Verkehrsaufkommen in einer Hauptstraße oder den Verlauf der letzten Grippeepidemie heranzuziehen. Auch das in seinen Grundzügen aus den 1970er Jahren stammende Datenschutzrecht orientiert sich an derartigen Lösungsmechanismen, indem es die Grundsätze der Erforderlichkeit und der Zweckbindung zu den entscheidenden Maßstäben für die Zulässigkeit der Verarbeitung personenbezogener Daten erklärte.

Heute richtet sich das Interesse immer stärker auf Big Data-Algorithmen, die sich nicht deterministisch an einer Aufgabe orientieren. Bei ihnen stehen Korrelationen, also statistische Zusammenhänge im Mittelpunkt, aus denen allerdings vielfach individualisierte Schlussfolgerungen gezogen werden. Die meisten kommerziell erfolgreichen Internetangebote verwenden solche Big Data-Verfahren, um zielgerichtete, personalisierte Werbung ohne die bei den klassischen Massenmedien unvermeidlichen hohen Streuverluste auszuliefern. Big Data-Algorithmen liefern auch die Grundlage für immer mehr Entscheidungen, die für unser Leben von existenzieller Bedeutung sind: ob wir einen Kredit erhalten und wenn ja zu welchen Konditionen, ob wir in eine Versicherung aufgenommen werden und wie viel Prämie wir dafür zu zahlen haben. Anstelle einheitlicher Preisangaben treten individualisierte Angebote, die sich an der (vermuteten) finanziellen Leistungsfähigkeit und Zahlungsbereitschaft orientieren. So müssen Apple-Nutzer aufgrund ihrer durchschnittlich höheren Zahlungsfähigkeit damit rechnen, für ihre Internetbestellungen mehr zu zahlen als die Verwender anderer Computersysteme.[8] Algorithmen schlagen vor, welcher Bewerber zu einem Vorstellungsgespräch eingeladen werden soll, wer für eine Beförderung infrage kommt und wer ein Entlassungskandidat ist. Auch im Gesundheitswesen nimmt die Bedeutung von Big

8 *Wilson*, If you use a Mac or an Android, e-commerce sites may be charging you more, Washington Post, 3.11.2014, abrufbar im Internet unter https://www.washingtonpost.com/posteverything/wp/2014/11/03/if-you-use-a-mac-or-an-android-e-commerce-sites-may-be-charging-you-more/ (letzter Zugriff 10.10.2015).

Data zu: Die Verknüpfung einer Vielzahl von Vitaldaten ermöglicht die frühzeitige Erkennung von Krankheiten. Anfragen bei der Internetsuche geben Hinweise auf Epidemien und ermöglichen schnelle Gegenmaßnahmen zu ihrer Eindämmung.

Schon jetzt liefern Big Data-Algorithmen nicht nur Datengrundlagen für menschliche Entscheider. Vielmehr entscheiden Computer anhand von Algorithmen zunehmend selbst, jedenfalls dort, wo die Sache klar scheint und wo die Entscheidung zeitkritisch ist: Etwa bei der Festlegung der Zahlungsmethode im Internethandel (Bestellung auf Rechnung oder nur gegen Vorauskasse), an der Supermarktkasse (Autorisierung der Zahlung per EC-Karte durch PIN oder Unterschrift), bei der Einsortierung eines Anrufers in die Warteschleife bei Telefonservices („gute" und „schlechte" Kunden werden anhand ihrer Telefonnummer identifiziert – die schlechten müssen länger warten, bis sie mit einem Berater im Call-Center verbunden werden). Immer mehr Anwendungen bestimmen bereits heute über unser Wohl und Wehe, nicht nur in den heftig diskutierten „autonom fahrenden" Kraftfahrzeugen, die dauernd untereinander, mit Werkstätten, dem Hersteller oder mit Navigationssystemen kommunizieren.

Ein zentrales Merkmal der algorithmischen Steuerung ist die Klassifizierung, d.h. die Zuordnung von Datenelementen zu bestimmten Gruppen. Diese erfolgt im Regelfall mittels – an sich nicht neuer – statistischer Verfahren. Versicherungen arbeiten seit eh und je mit Wahrscheinlichkeiten, die sich etwa in Sterbetafeln niederschlagen oder bei der Berechnung des Schadensfreiheitsrabatts in der Kfz-Versicherung. Schon vor langer Zeit wurden statistische Zusammenhänge zwischen dem Wohnort eines Kreditnehmers und seiner Zahlungskraft festgestellt und bei der Bewertung des Kreditrisikos berücksichtigt. Die Urform dieses heute als Georeferenzierung bezeichneten Verfahrens wurde in den 1920er Jahren in den USA eingeführt, als Stadtteile mit Bonitätsnoten versehen wurden. Auf Landkarten wurden die Viertel mit einer weniger zahlungskräftigen Bevölkerung rot umrandet, während die besseren Stadtteile gelb oder blau gekennzeichnet wurden. Die Bewohner rot markierter Stadtteile hatten praktisch keine Chance auf einen Kredit. Die Ergebnisse des „Redlining" sind bis heute zu erkennen: Da die Bewohner der entsprechend gekennzeichneten Stadtteile als nicht kreditwürdig galten, zogen die Familien, die es sich irgendwie leisten konnten, in besser klassifizierte Gegenden um, sodass sich die soziale Zusammensetzung der verbliebenen Bewohnerschaft immer weiter verschlechterte. Die finanzielle Risikobewertung führte auf diese Weise zur Verslumung ganzer Stadtregionen, wobei insbesondere ethnische Minderheiten auf der Strecke blieben.

Heute bedient man sich angesichts einer verbesserten Datenlage und leistungsfähiger Computer sehr viel differenzierterer Methoden zur Berechnung individueller Risiken. In die persönliche Bonitätsnote, den „Scorewert", fließen verschie-

denste Daten eines Menschen ein, die mit den Daten anderer Personen verglichen werden. Trotz dieser Individualisierung handelt es sich aber weiterhin um Klassifikationen, die im Unterschied zum Redlining nicht nur auf einem, sondern auf mehreren Faktoren beruhen. Der Scorewert ist eine Art Kopfnote des Kreditnehmers, die aber nicht sein tatsächliches, feststellbares Verhalten abbildet. Letztlich wird die individuelle Kreditwürdigkeit durch Vergleich von Personen berechnet, deren Daten hinsichtlich verschiedener Faktoren (etwa Wohnort, Alter, Geschlecht, Dauer eines Beschäftigungsverhältnisses, Anzahl der Bankkonten und der Handy-Verträge) denen des Kreditnehmers entsprechen. Auf diese Weise erhalten auch solche Personen ggf. eine schlechte Bonitätsnote, die bisher alle Kredite ordnungsgemäß zurückgezahlt haben und in einem festen Arbeitsverhältnis stehen, also auf den ersten Blick als kreditwürdig erscheinen. Allein die auf Wahrscheinlichkeiten beruhende Klassifikation führt zur Abwertung.

Der auf der Grundlage eines mathematisch-statistischen Verfahrens errechnete Scorewert soll Banken, Versandhändlern, Telekommunikationsunternehmen und Vermietern die Wahrscheinlichkeit des künftigen individuellen Zahlungsverhaltens offenbaren. Schlechte Bonitätsnoten führen in aller Regel dazu, dass der Betroffene keinen Kredit erhält oder mehr Zinsen zahlen muss oder dass eine Versicherung keinen Vertrag mit ihm abschließen will.

Auch in anderen Bereichen werden Algorithmen zur Mustererkennung eingesetzt. Intelligente Videosysteme, die Aufnahmen aus digitalen Überwachungskameras analysieren, verarbeiten die erfassten Bilder, um Geschlecht und Alter der aufgenommenen Personen zu klassifizieren und ihr Verhalten vorherzusagen. Diese Erkenntnisse werden für maßgeschneiderte Dienstleistungen oder aber zum Erkennen verwendet (Ladendiebe, Terroristen). Die Daten werden ggf. mit anderen verknüpft, die die Identifizierung des Einzelnen ermöglichen (z.B. Daten aus Treueprogrammen oder aus Smartphones). Auch ohne Identifizierung schafft die Videoanalyse die Voraussetzungen einer Klassifikation nach Alter, Hautfarbe und Geschlecht und nach anderen sichtbaren Merkmalen (z.B. von Senioren, Frauen in der Altersgruppe 20–30 Jahre etc.). Soweit die Videoanalyse nicht den Zweck verfolgt, Personen namentlich zu identifizieren, können die den jeweiligen Gruppen zuzurechnenden Individuen entsprechend den ihnen zugeteilten Eigenschaften oder Vorhersagen unterschiedlich behandelt werden. Dies führt zu Risiken der Diskriminierung und Stigmatisierung, etwa aufgrund des Geschlechts oder des ethnischen Hintergrunds.

Derartige Verfahren laufen ganz überwiegend heimlich ohne Wissen der Betroffenen ab. Vielfach merkt der Betroffene nicht einmal, dass er gerade Gegenstand einer automatisierten Bewertung ist: Der Besucher eines Einkaufszentrums bemerkt zwar vielleicht die Videokamera, die ihn überwacht, er kann aber nicht

erkennen, was mit den Aufnahmen im Hintergrund passiert. Auch bei der Internetnutzung erfolgt die individuelle Klassifizierung im Hintergrund: Welche Werbebotschaften dem Leser einer digitalen Zeitungsseite im Web präsentiert werden, ist häufig das Ergebnis eines komplexen Aushandlungsprozesses, bei dem – basierend auf dem individuellen Nutzerprofil – verschiedene Werbeanbieter ihre jeweiligen Gebote abgeben. Bei dieser Versteigerung kommt die Werbebotschaft zum Zuge, für die der Anbieter das höchste Gebot abgegeben hat.

Schließlich bedienen sich auch Sicherheitsbehörden vergleichbarer Algorithmen, um verdächtiges vom unverdächtigen Verhalten zu unterscheiden und risikoträchtige Personen zu identifizieren. In den USA werden bestimmte, als potentiell gefährlich klassifizierte Fluggäste gründlicher überprüft oder gänzlich abgewiesen. Bei der automatisierten massenhaften Überwachung des Internetverkehrs werden verdächtige Kommunikationsvorgänge selektiert und entsprechende Kommunikationsinhalte längerfristig gespeichert und gesondert ausgewertet. Voraussetzung für derartige Klassifikationen ist dabei stets, dass prinzipiell jede Interaktion erfasst und auf Muster untersucht wird, die bei der Bewertung herangezogen werden. Von Big Data zum Generalverdacht führt also ein ziemlich kurzer Weg.

IV. Algorithmische Diskriminierung

Besonders problematisch ist die algorithmische Klassifikation, wenn sie diskriminierende Folgen hat. Der Schutz vor Diskriminierung gehört zu den grundlegenden Menschenrechten (Art. 2 AERK). Niemand darf aufgrund von Rasse, Hautfarbe, des Geschlechts, der Sprache, der Religion, politischer oder sonstiger Ansichten, nationaler oder sozialer Herkunft, Vermögen, Geburt oder sonstigem Stand benachteiligt werden. Auch Art. 3 GG verbietet die Benachteiligung oder Bevorzugung eines Menschen „wegen seines Geschlechtes, seiner Abstammung, seiner Rasse, seiner Sprache, seiner Heimat und Herkunft, seines Glaubens, seiner religiösen oder politischen Anschauungen“. Das Allgemeine Gleichbehandlungsgesetz (AGG), das in Deutschland die europäische Antidiskriminierungsrichtlinie (RL 2000/78/EG) umsetzt, verbietet staatlichen Stellen und privaten Akteuren die unmittelbare oder mittelbare Diskriminierung, etwa bei der Begründung und beim Vollzug eines Beschäftigungsverhältnisses oder beim Zugang zu Bildung oder zu Gütern und Dienstleistungen einschließlich Wohnraum. Strenge Antidiskriminierungsregelungen gelten auch in den USA und vielen anderen Staaten.

Allerdings stellt nicht jede unterschiedliche Behandlung eine Diskriminierung dar. So heißt es etwa in § 20 AGG: „Eine Verletzung des Benachteiligungsverbots ist nicht gegeben, wenn für eine unterschiedliche Behandlung wegen der Religion, einer Behinderung, des Alters, der sexuellen Identität oder des Geschlechts ein sachlicher Grund vorliegt." Die Frage ist allerdings, was unter einem „sachlichen Grund" zu verstehen ist. Es liegt auf der Hand, dass bei einer einzelfallbezogenen Betrachtung nachvollziehbare, auf das konkrete Individuum bezogene Gründe für eine unterschiedliche Behandlung nachzuweisen sind. Dagegen liefern Methoden, wie sie bei Big Data-Analysen zum Einsatz kommen, nur anscheinend objektive Gründe für die Ungleichbehandlung, bei denen es sich bei näherer Betrachtung allerdings um nichts anderes als um Schlussfolgerungen aus Wahrscheinlichkeiten handelt. Wenn eine Person nicht nach ihrem tatsächlichen Verhalten, ihren Fähigkeiten und Eigenschaften, sondern nur gemäß einer mehr oder minder groben Klassifikation beurteilt wird, ist das Ergebnis zwangsläufig kein gerechtes Urteil, sondern eine besondere Form des Vorurteils. Mit einem Vorurteil haben derartige, mittels Scoring begründete Schlussfolgerungen gemein, dass sie der Betroffene kaum widerlegen kann. Anders als bei der vordergründigen Diskriminierung, etwa durch den Türsteher, der Angehörige mit bestimmten Merkmalen nicht durchlässt, erscheinen algorithmische Entscheidungen rational begründet und insofern objektiv. Dass letztlich auch bei solchen datenbasierten Entscheidungen dieselben Personen(gruppen) außen vor bleiben wie beim Türsteherbeispiel, wird dadurch verdeckt.

Die algorithmengesteuerte Klassifikation von Personen anhand statistischer Zusammenhänge hat vielfach diskriminierende Wirkungen. Sie sind den verwendeten mathematischen Verfahren inhärent und nicht etwa alleiniges Ergebnis inkorrekter Daten oder von Fehlfunktionen, auch wenn derartige „Fehler" angesichts des ungeheuren Datenumfangs und der vielfach nicht qualitätsgesicherten Daten kaum zu vermeiden sind mit der Folge, dass ein Vorgang bzw. Betroffener „falsch" klassifiziert wird.

In den USA werden bestimmte Vornamen unterschiedlich häufig von Menschen verschiedener Hautfarben gewählt.[9] Wie die Harvard-Forscherin und heutige Chefin der Technologieabteilung der Federal Trade Commission Latanya Sweeney beschreibt,[10] hat die systematische Auswertung von als Suchbegriffe

9 *Fryer/Levitt*, The Causes and consequences of distinctively black names, Quarterly Journal of Economics, S. 767, abrufbar im Internet unter http://pricetheory.uchicago.edu/levitt/Papers/FryerLevitt2004.pdf (letzter Zugriff 11.10.2015).

10 *Sweeney*, Discrimination in Online Ad Delivery, 2013.

eingegebenen Namen diskriminierende Auswirkungen: Die Google-Suche nach einem Namen, der eher auf einen Afroamerikaner hinweist, führt zur Einblendung von Informationen über vermeintliche Verbindungen zur Kriminalität, Vorstrafen oder Gefängnisaufenthalte, und zwar auch dann, wenn derartige Verwicklungen in dem konkreten Fall nicht vorliegen. Dagegen führt die Suche nach nicht typischerweise vermehrt von Afroamerikanern verwendeten Vornamen nicht zur Anzeige derartiger diskriminierender Sachverhalte. Im Hinblick darauf, dass heute praktisch sämtliche Bewerber um einen Job zunächst vom potentiellen Arbeitgeber oder privaten Jobvermittler gegoogelt werden,[11] ist naheliegend, dass diese Assoziationen die Chancen mancher Bewerber mindern.

Als sicher gilt, dass entsprechende Verfahren auch in anderen Bereichen Anwendung finden. So berichtet etwa Sandvig, dass Facebook-Nutzer mit Namen, die auf die Zugehörigkeit zu bestimmten Bevölkerungsgruppen hindeuten, von dem Unternehmen überdurchschnittlich häufig beschuldigt wurden, bei der Einrichtung Ihres Facebook-Kontos verbotenerweise einen erfundenen Namen zu verwenden.[12]

Problematisch wirken sich auch Bewertungssysteme aus, die allein oder überwiegend auf Informationen aus sozialen Netzwerken und anderen Internetforen beruhen. Verfahren zur Beurteilung etwa des persönlichen Einflusses,[13] der Leistungsfähigkeit und Zuverlässigkeit eines Job-Kandidaten oder der Kreditwürdigkeit beinhalten ein hohes Diskriminierungspotenzial. Je nach Präferenz des Betreibers eines solchen Dienstes können die ermittelten Scorewerte alle möglichen Informationen einfließen, etwa zum Alter, zur Gewerkschaftszugehörigkeit, zur politischen Einstellung, zur sexuellen Ausrichtung, zum Gesundheitszustand, zur Religionszugehörigkeit oder zum ethnischen Hintergrund der bewerteten Person. Es ist anzunehmen, dass potentielle Arbeitgeber auf diesem Weg an Informationen gelangen, die sie ansonsten legal nicht erhalten würden, beispielsweise im Hinblick auf eine mögliche Schwangerschaft einer Bewerberin.

11 Vgl. *Boyd/Levy/Marwick*, The Networked Nature of Algorithmic Discrimination, 2014, S. 55, abrufbar im Internet unter www.danah.org/papers/2014/DataDiscrimination.pdf (letzter Zugriff 12.10.2015).

12 *Sandvig*, Algorithms and Accountability – Vortrag auf der NYU Law Conference, abrufbar im Internet unter http://www.law.nyu.edu/centers/ili/algorithmsconference (letzter Zugriff 12.10.2015).

13 Vgl. *Gaffney/Puschmann*, Game or measurement? Algorithmic transparency and the Klout score, 2015, abrufbar im Internet unter http://www.researchgate.net/publication/276974372_Game_or_measurement_Algorithmic_transparency_and_the_Klout_score (letzter Zugriff 12.10.2015).

V. Algorithmentransparenz

Angesichts der zunehmenden Bedeutung von Algorithmen, die zur Klassifikation und Bewertung von Menschen herangezogen werden, wird die Frage nach ihrer Funktionsweise akuter. Nur wenn transparent ist, welche Daten in die jeweiligen Auswertungen und Bewertungsprozesse einfließen, nach welchen Kriterien die Klassifikation erfolgt und wie sie Entscheidungen beeinflussen, lassen sich Aussagen zu ihrer Rechtmäßigkeit und ethischen Vertretbarkeit treffen.

Das Datenschutzrecht trägt dieser Anforderung bislang nur unzureichend Rechnung. Die individuellen Auskunftsrechte der EG-Datenschutzrichtlinie von 1995 (Art. 12) und des deutschen Bundesdatenschutzgesetzes (§§ 6, 19, 34) beschränkten sich zunächst auf die zur Person des Betroffenen gespeicherte personenbezogene Daten. Dem Betroffenen ist auf Antrag Auskunft über die zu seiner Person gespeicherten Daten zu erteilen, auch soweit sie sich auf die Herkunft dieser Daten beziehen, die Empfänger oder Kategorien von Empfängern, an die die Daten weitergegeben werden, und den Zweck der Speicherung. Auch die Benachrichtigungspflichten bei bestimmten Verarbeitungsvorgängen, etwa bei der Speicherung ohne Kenntnis des Betroffenen (§§ 19a, 33 BDSG) beschränkten sich auf diese Angaben.

Das Bundesverfassungsgericht hatte bereits in seinem Volkszählungsurteil 1983 ausgeführt, dass die Wahrnehmung des Rechts auf informationelle Selbstbestimmung die Kenntnis des Betroffenen darüber voraussetzt, welche Daten über ihn gespeichert sind und wohin sie übermittelt werden.[14] Immer deutlicher wird allerdings, dass diese auf die jeweiligen Daten beschränkten Rechte zwar eine notwendige, nicht jedoch eine hinreichende Bedingung für die Transparenz der Datenverarbeitung sind. Bei der algorithmischen Bewertung werden vielfach Daten aus den unterschiedlichsten Quellen in Echtzeit zusammengeführt und daraus ein Scorewert erzeugt.

Immerhin hat der deutsche Gesetzgeber 2009 einen neuen § 28b BDSG eingefügt, der das Scoring regelt:

> „Zum Zweck der Entscheidung über die Begründung, Durchführung oder Beendigung eines Vertragsverhältnisses mit dem Betroffenen darf ein Wahrscheinlichkeitswert für ein bestimmtes zukünftiges Verhalten des Betroffenen erhoben oder verwendet werden, wenn

14 BVerfGE 65, 1, 43.

1. die zur Berechnung des Wahrscheinlichkeitswerts genutzten Daten unter Zugrundelegung eines wissenschaftlich anerkannten mathematisch-statistischen Verfahrens nachweisbar für die Berechnung der Wahrscheinlichkeit des bestimmten Verhaltens erheblich sind,
2. im Fall der Berechnung des Wahrscheinlichkeitswerts durch eine Auskunftei die Voraussetzungen für eine Übermittlung der genutzten Daten nach § 29 und in allen anderen Fällen die Voraussetzungen einer zulässigen Nutzung der Daten nach § 28 vorliegen,
3. für die Berechnung des Wahrscheinlichkeitswerts nicht ausschließlich Anschriftendaten genutzt werden
4. im Fall der Nutzung von Anschriftendaten der Betroffene vor Berechnung des Wahrscheinlichkeitswerts über die vorgesehene Nutzung dieser Daten unterrichtet worden ist; die Unterrichtung ist zu dokumentieren."

Zugleich wurde die für nicht-öffentliche Stellen einschlägige Auskunftsregelung um eine Sondervorschrift zur Auskunft über Scoring-Verfahren erweitert. § 34 Abs. 2 BDSG lautet:

„Im Fall des § 28b hat die für die Entscheidung verantwortliche Stelle dem Betroffenen auf Verlangen Auskunft zu erteilen über
1. die innerhalb der letzten sechs Monate vor dem Zugang des Auskunftsverlangens erhobenen oder erstmalig gespeicherten Wahrscheinlichkeitswerte,
2. die zur Berechnung der Wahrscheinlichkeitswerte genutzten Datenarten und
3. das Zustandekommen und die Bedeutung der Wahrscheinlichkeitswerte einzelfallbezogen und nachvollziehbar in allgemein verständlicher Form. [...]"

Der gute Wille, bei wahrscheinlichkeitsbasierten Entscheidungen für mehr Durchblick zu sorgen, ist unverkennbar. Trotzdem sind diese Vorschriften nicht mehr als ein erster Schritt zur Algorithmentransparenz. Zum einen beschränkt sich der Anwendungsbereich auf Verfahren, die bei der „Entscheidung über die Begründung, Durchführung oder Beendigung eines Vertragsverhältnisses" eingesetzt werden, in allen anderen Fällen, etwa beim Behavioral Targeting, besteht kein Anspruch. Zum anderen bleibt es bei einer Auskunftsregelung; die Betroffenen müssen selbst aktiv werden, um mehr Informationen zu erhalten. Schließlich beschränkt sich die Auskunftspflicht auf die zur Berechnung der Wahrscheinlichkeitswerte genutzten Datenarten und „das Zustandekommen und die Bedeutung

der Wahrscheinlichkeitswerte einzelfallbezogen und nachvollziehbar in allgemein verständlicher Form". Die Einzelheiten der Scoreberechnung müssen nicht mitgeteilt werden.

Der Bundesgerichtshof lehnte ein Auskunftsbegehren hinsichtlich der Faktoren, die zur Versagung eines beantragten Kredits geführt hatten, Anfang 2014 ab.[15] Zwar hätten die Betroffenen ein Recht zu erfahren, welche personenbezogenen, insbesondere kreditrelevanten Daten bei der Kreditauskunftei gespeichert sind, nicht jedoch die Formel, mit der der Scorewert berechnet wird. Die Berechnungsmethode sei ein Geschäftsgeheimnis der Auskunftei. Das BGH-Urteil zeigt, wie wenig Transparenz es bei algorithmenbasierten Entscheidungen immer noch gibt. Zwar müssen die Kreditauskunfteien inzwischen einmal jährlich auf Anfrage einen kostenlosen „Kontoauszug" zu den gespeicherten Daten übermitteln. Auch haben sie über wesentliche Faktoren zu informieren, die in Scorewerte eingeflossen sind. Die Scoreformel selbst aber bleibt geheim. Anders als beim legendären Coca-Cola-Rezept haben die Betroffenen keine Chance, sich entsprechenden Entscheidungen zu entziehen und wie bei der Cola etwa zu Pepsi oder Bionade zu wechseln. Ob bei der Commerzbank oder bei der Sparkasse, ob Vodafone oder Telekom: Wer einen Kredit beantragt oder einen Mobilfunkvertrag abschließen will, kann dem Scoring nicht entgehen.

Eine andere bereits in der EG-Datenschutzrichtlinie von 1995 enthaltene Vorschrift könnte für algorithmenbasierte Entscheidungen an Bedeutung gewinnen: das „Verbot automatisierter Einzelentscheidungen". Art. 15 verpflichtet die Mitgliedstaaten, jeder Person das Recht einzuräumen, „keiner für sie rechtliche Folgen nach sich ziehenden und keiner sie erheblich beeinträchtigenden Entscheidung unterworfen zu werden, die ausschließlich aufgrund einer automatisierten Verarbeitung von Daten zum Zwecke der Bewertung einzelner Aspekte ihrer Person ergeht, wie beispielsweise ihrer beruflichen Leistungsfähigkeit, ihrer Kreditwürdigkeit, ihrer Zuverlässigkeit oder ihres Verhaltens". Unbeschadet dieses grundsätzlichen Verbots kann eine solche Entscheidungen aber zulässig sein, wenn sie „im Rahmen des Abschlusses oder der Erfüllung eines Vertrags ergeht und dem Ersuchen der betroffenen Person auf Abschluss oder Erfüllung des Vertrags stattgegeben wurde oder die Wahrung ihrer berechtigten Interessen durch geeignete Maßnahmen – beispielsweise die Möglichkeit, ihren Standpunkt geltend zu machen – garantiert wird". Die Vorschrift des § 6a BDSG, der diese europarechtliche Vorgabe umsetzt, ist 2009 um einen neuen Abs. 3 ergänzt worden, der das Recht des Betroffenen auf Auskunft nach den §§ 19 und 34 auf den „logischen Aufbau der automatisierten Verarbeitung der ihn betreffenden Daten" erstreckt.

15 BGH, Urteil vom 28.1.2014 – VI ZR 156/13.

So sinnvoll diese Vorgaben auf den ersten Blick erscheinen, so gering ist bislang ihre praktische Relevanz. Sie beschränken sich auf Entscheidungen, die *ausschließlich* automatisiert erfolgen. Hinsichtlich der Entscheidungen, bei denen an irgendeiner Stelle ein Mensch beteiligt ist, gelten sie nicht. Darüber hinaus betreffen sie nur solche Einzelentscheidungen, die zur Verweigerung eines Vertragsabschlusses oder zu sonstigen erheblichen Beeinträchtigungen führen – vielfach ist aber überhaupt nicht absehbar, welche Konsequenzen sich ergeben, etwa bei dem oben genannten Beispiel rassisch differenzierter Suchergebnisse. Ferner ist die im deutschen Recht seit 2009 verankerte Transparenz in derartigen Fällen nur auf Nachfrage des Betroffenen im Zuge der Auskunftserteilung herzustellen und sie beschränkt sich auf den Einzelfall.

Die EU-Datenschutzreform könnte diese Defizite beheben. So hat das Europäische Parlament in seiner Entschließung zur Datenschutzgrundverordnung die Einführung einer Informationspflicht über das Profiling vorgeschlagen.[16] Die Betroffenen sollten danach Informationen über „[...] Angaben über das Vorhandensein eines Profilings, auf Profiling gestützte Maßnahmen und die beabsichtigten Auswirkungen des Profilings auf die betroffene Person“ und „aussagekräftige Informationen über die Logik einer automatisierten Datenverarbeitung“ erhalten. Der am 6. Dezember 2015 zwischen den EU-Gremien im Rahmen des Trilogs erzielte Kompromiss übernimmt diese Forderung in Art. 14a Nr. 2 Buchst. f der Datenschutzgrundverordnung.[17]

Auch wenn damit keine vollständige Algorithmentransparenz hergestellt wird, handelt es sich um einen ersten wichtigen Schritt in diese Richtung, bei dem es hoffentlich nicht bleiben wird.

16 Legislative Entschließung des Europäischen Parlaments vom 12.3.2015 zu dem Vorschlag für eine Verordnung zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr, abrufbar im Internet unter http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2014-0212+0+DOC+XML+V0//DE (letzter Zugriff 13.10.2015).

17 European Parliament, Consolidated text of the data protection regulation, abrufbar im Internet unter http://ow.ly/d/47uW (letzter Zugriff 1.1.2016).

Das europäische Datenschutzrecht befindet sich nach der jüngst verabschiedeten Datenschutz-Grundverordnung in einer Zeit des Umbruchs. Ein Generationenwechsel steht unmittelbar bevor. Auch das EuGH hat mit seinen Entscheidungen zu Safe Harbor und Google Spain wesentlich zu den aktuellen Entwicklungen beigetragen.

Das diesjährige Jahrbuch für Informationsfreiheit und Informationsrecht greift den frischen Wind in der Datenschutzdebatte auf und präsentiert bedeutsame Inhalte der Neuerungen. Dabei wird nicht nur länderspezifischer Fortschritt unter die Lupe genommen. Auch globale Fragestellungen wie die Verhandlungen zwischen der EU und den USA über die Transatlantic Trade and Investment Partnership (TTIP) und die multilateralen Verhandlungen über das Trade in Services Agreement (TISA) finden Einzug in das Werk.

Die in diesem Band enthaltenen Beiträge vereinen wissenschaftliche Expertise mit praxistauglichen Hilfestellungen und verleihen dem Jahrbuch 2015 Themenvielfalt und Aktualität.

www.lexxion.de

€ 78,–
ISBN 978-3-86965-281-8

9 783869 652818

Der juristische Verlag
lexxion
Berlin
Brüssel