verordnung die selbstbestimmte Einwilligung einer betroffenen Person in den einzelnen EU-Mitgliedstaaten nicht mehr unterschiedlich stark oder gering gewährleistet werden.

Der Beitrag erschien in vorgänge. Zeitschrift für Bürgerrechte und Gesellschaftpolitik #211/212 (1/2018). Wir danken Redaktion und Autorin für die freundliche Genehmigung zum Nachdruck.

Referenzen

Arendt, Hannah 1967: Vita activa oder vom Tätigen Leben, München

Buchner, Benedikt 2013: Outsourcing in der Arztpraxis – zwischen Datenschutz und Schweigepflicht, M… S. 337–342

Borasio, Gian Domenico 2014: selbs…

Ernst, Stefan 2017: Die Einwilligung … nung, ZD (Zeitschrift für Datensch…

Fries, Jakob Friedrich 1803: Philosophische Rechtslehre und Kritik aller positiven Gesetzgebung, Jena

Guradze, Heinz 1956: Der Stand der Menschenrechte im Völkerrecht, Göttingen

Hoffmann-Riem, Wolfgang 2016: Innovation und Recht – Recht und Innovation, Frankfurt a. Main

Jarass, Hans D. 2016: GRCh, Charta der Europäischen Grundrechte, Kommentar (3. Auflage), München

Kühling, Jürgen/Buchner, Benedikt 2017: Datenschutz-Grundverordnung, Kommentar, München

Kafka, Franz 2004: Zur Frage der Gesetze und andere Schriften aus dem Nachlass (Taschenbuchausgabe), Frankfurt a. Main

Ramsey, Robert 1970: The Patient as Person, Yale University Press, New Haven

Schirrmacher, Frank 2011: Payback. Warum wir im Informationsalter gezwungen sind zu tun, was wir nicht tun wollen, und wie wir die Kontrolle über unser Denken zurückgewinnen, München

Schmale, Wolfgang/Tinnefeld, Marie-Theres 2017: Europa durch Menschenrechte, in: Datenschutz und Datensicherheit (DuD) Heft 6, S. 343–47

Spindler, Gerold 2012: Persönlichkeitsschutz im Internet – Anforderungen und Grenzen einer Regulierung, Gutachten F zum 69. Deutschen Juristentag, F 1 ff.

Tinnefeld, Marie-Theres 2018: Überleben in Freiräumen. 12 Denk-Stücke, Wien/Köln/ Weimar

Tinnefeld et al. 2017: Einführung in das Datenschutzrecht. Datenschutz und Informationsfreiheit in europäischer Sicht (6. Auflage), Berlin, Boston

Will, Rosemarie 2018: Kein Ende des jahrzehntelangen Rechtsstreites zum Erwerb eines Medikaments zur Selbsttötung in: Grundrechte-Report 2018

erschienen in der FIfF-Kommunikation, herausgegeben von FIfF e.V. - ISSN 0938-3476 www.fiff.de

… anasie"-Systems am Beispiel der Ermor… der „schwierigen" Kindern durch Aus… …en von Nazi-Psychiatern befasst sich der Film „Nebel im August" (2016, Regie: Kai Wessel) nach dem gleichnamigen, 2008 erschienenen Buch, in dem Robert Domes das Leben und die Ermordung von Ernst Lossa darstellt, der zum fahrenden Volk der Jenischen gehörte, einer heterogenen Gruppe, die von den Nationalsozialisten als „Zigeuner" bezeichnet wurde.

2 Internationaler Pakt über bürgerliche und politische Rechte (IPbpR).

3 Tinnefeld, in Tinnefeld et al. 2017: Prolog S. XIX.

4 IT-Grundrecht = Grundrecht auf Gewährleistung der Vertraulichkeit und Integrität informationstechnischer Systeme; s. BVerfG, Urteil des Ersten Senats v. 27. Februar 2008 – 1 BvR 370/07 (=BVerfGE 120, 274 ff.).

5 Buchner/Kühling, in Kühling/Buchner 2017: Art. 7 Rn. 5 ff.

6 Tinnefeld, in Tinnefeld et al. 2017: 208 ff.

7 „Sprache des Rechts. Vermitteln, Verstehen, Verwechseln." Interdisziplinäre Arbeitsgruppe an der BBAW, http://www.bbaw.de/iag/ag_sprache/ueber.html.

Rainer Rehak

(Meta-) Daten im Zeichen der Sicherheit?

Zum staatlichen Umgang mit vernetzten Datenbeständen

Mit der zunehmenden Digitalisierung hinterlassen die Handlungen von Menschen Metadaten in den jeweiligen Systemen. Sie werden für die kommerzielle Profilerstellung, aber auch für folgenschwere polizeiliche und geheimdienstliche Zwecke ausgewertet. Über diese technikgläubige Herangehensweise muss dringend diskutiert werden.

Seit ihrer Existenz sammeln und speichern staatliche Stellen Informationen über ihre BürgerInnen. Auch die damit eng verbundene Grenze zwischen als notwendig erachteter Verwaltung und weit darüber hinausgehenden Kontrollabsichten ist seit jeher Gegenstand gesellschaftlicher Diskussionen. Seit Jahrzehnten befinden wir uns nun im Prozess einer zunehmenden automatisierten Datenverarbeitung persönlicher, geschäftlicher sowie gesellschaftlicher Interaktionen – inzwischen lapidar *Digitalisierung* genannt. Informationen werden nicht mehr in dunklen Kellern in Form von papierenen Aktenmetern angelegt, aufbewahrt und mühsam manuell durchsucht, sondern können in vernetzen informationstechnischen Systemen erzeugt und verarbeitet werden. Volltextsuche, Mehrfachindexierung, Sortieren, Filtern und effizientes Speichern sind dabei nur noch Fingerübungen der Informatik, die in jedem Informatikbuch über Algorithmen nachzulesen sind.[1] Neuere Methoden drehen sich eher um das Finden von Korrelationen durch statistische Analysen oder das Entdecken ähnlicher Strukturen durch Muster(wieder)erkennung, beispielsweise mit *lernenden* künstlichen neuronalen Netzen (KNN) oder anderen heuristischen – und damit nicht-exakten – Ansätzen. Ist die Datengrundlage vergleichsweise groß und vielfältig, so fällt häufig der unscharfe, aber politisch und wirtschaftlich wirkmächtige Begriff *Big Data*. Wenn es darum geht, die Ergebnisse solcher Herangehensweisen zu interpretieren, stellen sich jedoch weitreichende Fragen: Was sagen beispielsweise Korrelationen über kausale Zusammenhänge aus?

Und was sind *ähnliche* Strukturen? Welche *Muster* können und sollen überhaupt erkannt werden?

Die „digitalisierte Gesellschaft"

Wenn wir über die *digitalisierte Gesellschaft* sprechen, so ist damit häufig gemeint, dass jegliche persönliche Informationsverarbeitung mittels Digitalcomputern geschieht. Wir reichen unsere Steuererklärung digital ein, tragen ein Mobiltelefon bei uns, nutzen digitale Plattformen zum Informationstausch und Warenkauf, verwenden E-Mails und Instant Messenger zur Kommunikation, haben unsere Backups in der ominösen *Cloud* und fragen den *Wahl-O-Mat* nach unseren Wahlpräferenzen. Doch ein Fokus allein auf die individuelle Nutzung greift zu kurz, denn permanent sind wir in staatlichen und wirtschaftlichen Prozessen von digitalen vernetzten Informationssystemen umgeben und Gegenstand ihrer Verarbeitung: von der Abwicklung des Flugverkehrs, den Rentenverwaltungssystemen oder dem zentralen Fahreignungsregister in Flensburg (früher Verkehrszentralregister) über die Polizeiverwaltungs-, Fahndungs- oder Fallbearbeitungssysteme bis hin zu Krankenkassen-Verwaltungsstrukturen, Mautsystemen, den Einwohnermelde- und Finanzämtern, der Schutzgemeinschaft für allgemeine Kreditsicherung (Schufa), Unternehmens- sowie Bankensystemen und schließlich dem *Internet der Dinge* oder den Sensorensystemen für das *autonome* Fahren.

Neben vielen interessanten Aspekten dieser allgemeinen Entwicklung, beispielsweise dass eine der weltweit größten Zimmervermietungen gar keine Zimmer besitzt (AirBnB) oder die weltweit größte Enzyklopädie von Freiwilligen befüllt und administriert wird (Wikipedia), gibt es spezielle Eigenschaften aktueller digitaler Systeme. Hier fallen nicht nur die direkt verarbeiteten Daten an wie beispielsweise die eigentlichen Kommunikationsinhalte, die abgegebene Steuererklärung, die Schlafrhythmusdaten von Fitness-Apps, die Rechnungen in PDF-Form, die *gekauften* (bzw. tatsächlich nur lizenzierten) E-Books und andere Medieninhalte, sondern auch eine ganze Menge der viel zitierten *Metadaten* an. Metadaten sind Daten, die die Umstände der Datenverarbeitung beschreiben, wie etwa BesitzerIn und Erstellungszeitpunkt eines Dokumentes, die involvierten Sender- und Empfängernummern eines Telefongesprächs bzw. Nachrichtenaustauschs, die aufgerufenen Unterseiten einer Webseite und deren Ansichtsdauer, der aktuell verbundene Funkmast von Mobiltelefonen, die IP-Adresse von WebseitenbesucherInnen oder die genauen Nutzungszeiten von Kommunikationsdiensten. Diese Metadaten entstehen zwar nicht notwendigerweise, aber die meisten Systeme sind so gebaut, dass nahezu alle Aktivitäten festgehalten – geloggt – werden. Dafür gibt es teilweise technische Gründe, eine effektivere Fehlersuche oder schnellere Angriffserkennung, aber es überwiegen tatsächlich eher kommerzielle Gründe, beispielsweise Abrechnungsprozesse oder die Möglichkeit einer detaillierten Benutzerprofilerstellung ebenso wie zum A/B-Testen von Alternativinhalten.

Das Konzept *Metadaten* ist zwar nicht neu, aber ein Eingangsstempel auf einem papierenen Brief lässt sich nicht so automatisiert und massenhaft auswerten wie die Digitalversion. Hinzu kommt der Umstand, dass mit den aktuellen Hardwarekosten und Softwaredesigns das Behalten und Speichern von Daten und Metadaten viel billiger und weniger aufwändig ist als das Löschen, weil Daten beispielsweise auch aufeinander verweisen und so auch in andere Kontexte hinein verknüpft sind.

Daten oder Metadaten

Gerade in der Politik wird immer noch oft die Ansicht vertreten, Metadaten seien weniger aussagekräftig und daher weniger schützenswert als Inhaltsdaten. Doch offenbaren beispielsweise Metadaten von Kommunikationsvorgängen – die sogenannten Verkehrsdaten – den kompletten sozialen Graphen. Der zeigt, wer von wo mit wem und wann kommuniziert. Daraus lässt sich direkt ableiten, welche Gruppen und Zusammenhänge es gibt und wer die zentralen, vernetzten Personen sind. Aus den Kommunikationszeiten wiederum lässt sich in der Regel auch die Art der Beziehung ablesen – beruflich oder privat, lose oder intim, stabil oder dynamisch. Anrufe bei Anwaltskanzleien für Arbeitsrecht, HIV-Hilfestellen, psychologischen Praxen, Escort-Services oder Familienplanungszentren brauchen kaum weitere Inhaltsdaten, um für sich bereits aussagekräftig zu sein. Darüber hinaus lassen sich Metadaten auch mit anderen Informationen verknüpfen, und so offenbaren die Ortsdaten nicht nur Essensgewohnheiten (in Verknüpfung mit Restaurant-Listen) oder Gesundheitszustand (Arzt- oder Krankenhausverzeichnisse), sondern auch komplette Verhaltensprofile.[2] Denn würden die Metadaten (hier konkret Orts- und Zeitangaben) einer hypothetischen Vorratsdatenspeicherung[3] beispielsweise mit Informationen zu angemeldeten Demonstrationen oder anderen politischen Aktionen verknüpft, so erhielte etwa eine Behörde komfortabel nahezu vollständige Anwesenheitslisten dieser Veranstaltungen. Metadaten sind also ebenso aussagekräftig wie Inhaltsdaten, sie müssen nur anders ausgewertet werden. Auch deswegen schützt Artikel 10 des Grundgesetzes nicht nur die Kommunikation, sondern auch deren Umstände, wie die zugehörigen Metadaten. Die digitalen Spuren all unserer Handlungen liegen also entweder direkt von uns produziert (beispielsweise Kommunikationsdaten) oder indirekt erzeugt (Kunden-, Steuer-, Maut-, Kranken- oder Rentendaten) vor. Einzig rechtliche Einhegungen sorgen dafür, dass nicht einfach alle Daten zusammengeführt werden, um ein umfassendes digitales Abbild der menschlichen Welt zu erschaffen.[4] Im Allgemeinen gilt im europäischen Datenschutzrecht daher ein sogenanntes Erhebungsverbot mit Erlaubnisvorbehalt für die Verarbeitung personenbezogener Daten, demzufolge grundsätzlich nichts gespeichert werden darf, es sei denn es gibt explizite Gründe für eine Erhebung, z. B. eine informierte Einwilligung.[5]

Polizeidateien

In Deutschland legen Polizeien eigene Datenbanken über Personen an, die sie als relevant erachten. Diese sogenannten *Dateien* benötigen je nach Bundesland manchmal eine Errichtungsanordnung, für den Bund jedoch immer.[6] Sie definieren den Zweck, den betroffenen Personenkreis, Datenquellen, Regeln der Datenübermittlung an andere Stellen oder Höchstspeicher- und Prüffristen. In Berlin muss beispielsweise bei jeder neuen *Datei* der Berliner Beauftragte für Datenschutz und Informationsfreiheit unterrichtet werden, in Hamburg nur dann, wenn die Errichtung mit „besonderen [...] Problemen" verbunden ist.[7]

Eine solche Regelung aber sorgt bisweilen für merkwürdige Situationen, wie in Hamburg im Jahre 2014, wo das dortige Landeskriminalamt (LKA) eine Transparenzgesetz-Anfrage bezüglich der Existenz einer Sportgewalt-Datei verneinte. Dies entsprach „bedauerlicherweise" schon neun Jahre lang nicht mehr der Wahrheit, wie sich später herausstellte.[8] Die verschiedenen Polizeien erstellen so also jeweils eigene Dateien mit jeweils eigenen Formaten, für die jeweils die eigenen Landesinnenressorts verantwortlich sind. Dieses Unorganisiertheit ist im Sinne der vertikalen Gewaltenteilung beabsichtigt und wirkt gewissermaßen machtbeschränkend.

Interessant wird es dann, wenn federführend durch das Bundesministerium des Innern (BMI) und praktisch ausgeführt vom Bundeskriminalamt (BKA) Dateien bundesweit angelegt werden, die *INPOL-Verbunddateien*. Auch sie sind zweckgebunden und sollen Angaben zu den als interessant angesehenen Personen enthalten. Befüllt bzw. verwendet werden diese rund 150 Verbunddateien aber von den Landes- und Bundesbehörden gemeinsam, teilweise auch von Geheimdiensten. Das Konzept der Verbunddateien ist immer wieder Gegenstand von Kritik.[9] In einem prominenten Fall ging es um Teilnehmende einer Anti-Atom-Demonstration, deren Namen vor zwei Jahren vom Verfassungsschutz – laut BMI zu Recht – in eine gemeinsam mit dem BKA genutzte Projektdatei aufgenommen worden sind, mit der Begründung, Kernkraftkritik sei ja Systemkritik.[10] Für das BKA war eine so willkürlich gefüllte Datei jedoch praktisch nutzlos.

Ein weiteres Beispiel war die „Zentraldatei politisch-motivierte Kriminalität, links" (PMK-links Z), die im Jahre 2012 durch den damaligen Bundesdatenschutzbeauftragten Peter Schaar analysiert wurde. Im Gegensatz zu gemeinsamen Verbunddateien kann bei Zentraldateien nur das BKA schreibend zugreifen, es muss also selbst die Einträge prüfen. Schaar hatte dabei so viele Rechtsverstöße festgestellt, dass das BKA ca. 90 Prozent der Einträge löschen musste.[11] Die absolute Mehrheit der gespeicherten Personen waren also illegal in staatlichen Datenbanken gelandet, von deren Existenz sie nichts wussten und deren Auswirkungen auf etwaige Zuverlässigkeits-Überprüfungen komplett unklar waren. Die Konsequenzen einer solchen Speicherungspraxis zeigen sich besonders im Falle einer Rasterfahndung[12], denn der Abgleich mit einer Datei, die fast ausschließlich fälschlich gespeicherte Personendaten enthält, kann für die Betroffenen verheerende Folgen haben – für die eigentlich Gesuchten wiederum ist eine solche Praxis sehr von Vorteil. Hier offenbart sich das generelle Problem gemeinsam genutzter Datenbestände: Informationen werden von einem Akteur in einem Kontext mit einer bestimmte Absicht erhoben und dann – dekontextualisiert – als Daten gespeichert. Mit der Nutzung durch andere Akteure werden sie dann – meist ganz anders – rekontextualisiert. Dass aber die Kontexte der Erhebung und die der Nutzung zusammenpassen, muss akribisch sichergestellt werden, insbesondere wenn es sich bei den Akteuren um staatliche Stellen mit großer und/oder verdeckter Wirkmacht handelt. Die Konsequenzen einer solchen unzureichender Datenhaltung sind auch kürzlich wieder prominent zutage getreten. Die verweigerten 32 Akkreditierungen für JournalistInnen beim G20-Gipfel in Hamburg lassen sich mehrheitlich[13] auf falsche, nicht-aktualisierte oder schlicht illegal gespeicherte Daten solcher Verbunddateien zurückführen. Dabei haben die wenigsten Opfer solcher Datenpraktiken das Glück, in diesen publikumswirksamen Berufsfeldern zu arbeiten.

Die Liste derartig überbordernder Datenbanknutzung und falscher Einträge lässt sich nach wie vor millionenfach fortsetzen, sodass in den Medien nun von der „Spitze des Eisbergs" gesprochen wird[14] und BKA-Präsident Holger Münch sich an Generalrevisionsforderungen abarbeiten muss[15], während seine Kolleginnen und Kollegen in guter Verfassungsschutzmanier die Beweise des eigenen organisationalen Fehlverhaltens vernichten.[16]

Existenz fragwürdig, Prozesse intransparent, Daten veraltet

Es zeichnet sich ab, dass die schlechte Performanz solcher Datenbanken nicht die Ausnahme sondern die Regel darstellt. Eine sinnvolle Nutzung wäre rein theoretisch nur möglich durch mehr Qualitätssicherungspersonal, detailliertere Dateneingangsprüfungen, strikte Eintragsverbote beispielsweise für Personen mit Freisprüchen, kontextbeschreibende Annotationen der Daten, sinngebende Verweise auf Akten, Verfahren oder Hintergründe und regelmäßige, aufwändige Datenpflege inklusive restriktiver Löschfristen. Denn Daten veralten, verändern sich, müssen korrigiert oder gelöscht werden. Sollte das jedoch mit den vorhandenen Mitteln gar nicht möglich sein, so müssen Nutzen und Erforderlichkeit solcher Dateien generell infrage gestellt werden. Bei einer ständig unterbesetzten Polizei, die schon jetzt viele konventionelle Spuren kaum verfolgen kann, sind derartig komplex zu betreibende, löchrige, veraltete, illegale Datenbestände sogar schädlich. „Ganz klar: Unnötig gespeicherte Daten schaffen nicht mehr, sondern weniger Sicherheit", befand überraschend auch Justizminister Heiko Maas (SPD) im Kontext des G20-Akkreditierungsdebakels.[17] Ebenso klagen BKA-interne Analysten über zu viele irrelevante Daten in den Verbunddateien; insbesondere dort, wo Geheimdienste mit im Boot sind, da diese immer auf mehr Informationen aus sind, unabhängig davon, ob sie sich sauber überprüfen lassen.[18]

Lange Zeit, so scheint es, war die Nutzung solcher Dateien politisch gewollt. Auch diese Entwicklung muss im Kontext der Vernetzung und Digitalisierung sowie ihrer hehren Verheißungen verstanden werden: Auch hier spielen Technikgläubigkeit und mechanistische Weltbilder eine wesentliche Rolle, denn oft herrscht bezüglich der Kriminalitäts- und Terrorbekämpfung die Vorstellung einer Suche nach der „Nadel im Heuhaufen", wofür ja zuerst der ganze Heuhaufen benötigt würde.[19] In Deutschland werden bislang keine Big-Data-Analysen auf Basis polizeilicher Dateien durchgeführt und Datenbestände mit verschiedenen Zwecken (etwa des Staatsschutzes, der Bekämpfung der Organisierten oder der Wirtschaftskriminalität) dürfen auch nicht verkettet werden. Trotzdem sehen viele Personen in politischen Führungspositionen eine verheißungsvolle Zukunft in der Abkehr vom restriktiven Datenschutz hin zum Datenreichtum als Lösungsansatz für wirtschaftliche, ökologische oder polizeiliche Aufgabenstellungen.[20] Dieser Denkweise sind Trennungsgebot, Verkettungsverbot bzw. Zweckbindung ein Dorn im Auge.

Wilde Erfahrungen mit (Meta-) Daten

Was mit den angesammelten Daten passiert, wenn es zu wenige der oben beschriebenen Beschränkungen gibt, sehen wir beispielsweise in China, wo gerade ein *Sozialkredit*-Punktestand aller BürgerInnen aufgebaut wird. In dieser Datenbank wird gespei-

chert, wer bei Rot über die Ampel geht, wer Rechnungen nicht bezahlt oder wer sich kritisch über die Regierung äußert.[21] Ein anderes Beispiel ist die die verhängnisvolle Metadatennutzung für Drohnentötungen des US-Militärs in Pakistan oder Jemen. Auch der deutsche Bundesnachrichtendienst (BND) hat dafür Kommunikations-Metadaten sowie Stammdaten wie zugehörige Namen und Adressen beigesteuert.[22] Für solche Drohnenangriffe werden dann konkret nicht bekannte Personen anvisiert, weil bestimmte Muster passen. Bei diesen „signature strikes"[23] werden Eigenschaften und Zusammenhänge definiert, etwa regelmäßige Aufenthalte an bestimmten Orten, Telefonanrufe oder ähnliche Bewegungsmuster, wie sie andere, bereits bekannte Personen aufweisen. Diese Art von Datenverknüpfung wird allein mit Metadaten möglich, mit tödlichen Folgen für die Getroffenen.

Es gibt jedoch auch ganz andere Verwendungen von Metadaten, die keine komplexen Modelle brauchen, wie etwa die geheime Sammlung von Kompromat gegen „Gefährder" durch den US-amerikanischen Geheimdienst NSA zeigt. In einem der Fälle wurden massenhaft völlig legale, aber sozial brisante Webseitenzugriffe auf Pornographiewebseiten auf Vorrat gespeichert. Die damit erlangten Erotikvorlieben der Nutzer sollten dann verwendet werden, um die Zielpersonen bei Bedarf zu erpressen. In anderen Fällen wurden einfach alle BesucherInnen von Webseiten wie WikiLeaks (Enthüllungsplattform), TheTorProjekt.org (Anonymisierungssoftware) oder PirateBay (File-Sharing-Seite) auf Vorrat dokumentiert, vermutlich für eine spätere noch zu definierende Verwendung.[24] Hier wird erkennbar, welche Wirkung Metadaten entfalten können. Ähnlich datengetriebene Herangehensweisen sind bei der „prädiktiven Polizeiarbeit" in Teilen der USA erkennbar, bei der jedem Menschen mittlerweile *Gefahrenbewertungen* zugewiesen werden. Dieser *Gefährder-Score* berechnet sich nach den Datenmodellen der Hersteller, die zum Schutz von Betriebsgeheimnissen leider nicht nachvollzogen werden können.[25] Nur soviel ist bislang bekannt: Persönliche Daten aus Social-Media-Plattformen spielen genauso eine Rolle wie – bemerkenswert – der Kontakt zu Gewaltopfern.[26]

Gerade in Bezug auf Kommunikationsdaten ist auch in Deutschland eine starke Tendenz zur Datenanhäufung und -nutzung erkennbar. Die Zahlen von ermittlungsbezogenen Funkzellenabfragen nach Strafprozessordnung (StPO) nehmen stark zu, wobei großflächig und regelmäßig auf die vorhandenen Metadaten der Vorratsdatenspeicherung zurückgegriffen wird[27], ebenso wie die Nutzung von metadatenerzeugenden *Stillen SMS*. Interessant in diesem Zusammenhang: Im Jahre 2015 wurde die Firma Rola Security – Anbieter für polizeiliche Fallbearbeitungssoftware mit Überwachungsschnittstellen – von der Telekom gekauft.[28] Damit konnte die Telekom in Bezug auf Telekommunikationsüberwachung alles bequem aus einer Hand liefern.

Bei allen Anwendungsfällen fällt auf, dass die Betroffenen keine oder nur geringe Einfluss- und Beschwerdemacht haben, weil die Aktivitäten geheim ablaufen, weil ein späterer Widerspruch sinnlos wäre oder alles zusammen.

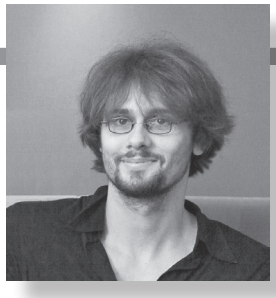
Big Data, künstliche Intelligenz und Technikgläubigkeit

Ganz allgemein gesprochen ist offensichtlich gerade staatlichen Akteuren weder bewusst, was automatisierte Datenauswertung kann bzw. nicht kann, noch was die Voraussetzungen dafür sind oder wie verheerend die Auswirkungen für die Betroffenen sein können.

Für die automatisierte Datenauswertung muss immer klar sein, was die gesuchten Zusammenhänge ausmacht, worin sie also genau bestehen. Mit den traditionellen informatischen Mitteln wie Suchen, Filtern, Sortieren sind immer auch formale Beschreibungen des Gesuchten notwendig. Es ist eben nicht möglich, Algorithmen auf eine Datensammlung anzusetzen und einfach nach *Terroristen* oder *Gefährdern* suchen zu lassen, denn wir haben bislang nicht einmal eine nicht-formale, allgemein anerkannte Definition von *Terrorismus* oder *Gefährderverhalten*. Wonach suchen wir also? Und gerade bei neuen Analyse- und Auswertungsmethoden mit bislang unklarer Wirkungsweise wie künstlichen neuralen Netzen muss genau abgewogen werden, was die Konsequenzen von Fehlanalysen sind, um den möglichen Nutzen damit abzugleichen.

Wenn beispielsweise der Google-Bilderkennungsalgorithmus ein Bild falsch klassifiziert, Amazon ein unpassendes Buch empfiehlt oder AlphaGo vielleicht auch einmal eine Partie verliert[29], ist die Konsequenz doch ungleich erträglicher als wenn fehlerhafte Rückfallvorhersagesoftware bei Gerichtsprozessen überwiegend Menschen dunkler Hautfarbe hinter Gitter bringt, JournalistInnen ihre Arbeit nicht mehr ausüben können, Menschen ihre politischen Aktivitäten einschränken, um keine verhängnisvollen Spuren mehr zu hinterlassen oder afghanische Bauersleute sterben, weil sie am falschen Ort Hochzeit gefeiert haben. Über diese Auswirkungen automatisierter Datenverarbeitung müssen wir dringend diskutieren, bevor wir eine Gesellschaft in – wenn auch manchmal nur ungewollt – ungerechte Technik gießen. Gerade auch Technikerinnen und Techniker müssen sich hier politisch zu Möglichkeiten und vor allem Grenzen von informationstechnischen Herangehensweisen äußern; oder um es sinngemäß mit dem Computerpionier und Gesellschaftskritiker Prof. Dr. Joseph Weizenbaum zu sagen: „Früher übergab man ein Problem dem Computer, wenn man es verstanden hatte. Heute ist es zunehmend anders herum." Diese Entwicklung gilt es zu stoppen.

Rainer Rehak



Rainer Rehak beschäftigt sich seit rund zehn Jahren mit dem Themenfeld *Informatik und Gesellschaft*. Er studierte Informatik und Philosophie in Berlin, Hong Kong und Peking. Während des Studiums arbeitete er am Lehrstuhl für I*nformatik in Bildung und Gesellschaft* von Wolfgang Coy. Aktuell promoviert er am Weizenbaum-Institut für die vernetzte Gesellschaft und lehrt in den Bereichen Datenschutz/Datensicherheit, sowie Informatik und Gesellschaft.

Dieser Text erschien zuvor in gekürzter Fassung in der Civil Liberties and Police (CILIP) 114 des Instituts für Bürgerrechte & öffentliche Sicherheit unter dem Titel „Die Datenschatten". Vielen Dank auch an Matthias Monroy und Heiner Busch für ihr wertvolles Feedback.

Anmerkungen

1 *Siehe z. B. Cormen, Leiserson und Rivest: Algorithmen – Eine Einführung, De Gruyter Oldenbourg, 2013*
2 *zeit.de v. 24.2.2011, http://www.zeit.de/digital/datenschutz/2011-02/vorratsdaten-malte-spitz*
3 *Siehe die damaligen Pläne für das Gesetz zur Einführung einer Speicherpflicht und Höchstspeicherfrist für Verkehrsdaten.*
4 *Vergleiche dazu die Situation in China: Deutschlandfunk-Kultur, Weltzeit v. 5.9.2017, http://www.deutschlandfunkkultur.de/chinas-sozialkredit-system-auf-dem-weg-in-die-it-diktatur.979.de.html?dram%3Aarticle_id=395126*
5 *Siehe das Volkszählungsurteil des Bundesverfassungsgerichts vom 15.12.1983, http://www.servat.unibe.ch/dfr/bv065001.html*
6 *Siehe § 34 BKA-Gesetz, § 490 Strafprozessordnung (StPO) oder § 49 ASOG Berlin*
7 *§ 26 Gesetz über die polizeiliche Datenverarbeitung Hamburg*
8 *Netzpolitik.org v. 17.1.2016, https://netzpolitik.org/2016/hamburger-polizei-fuehrt-geheime-datei-zur-sportgewalt-beluegt-buerger/*
9 *Zeit.de v. 24.9.2014, http://www.zeit.de/politik/deutschland/2014-09/bundeskriminalamt-daten-buerger-straftaeter*
10 *Deutschlandradio: Informationen am Morgen v. 2.9.2017, http://srv.deutschlandradio.de/themes/dradio/script/aod/index.html?audioMode=3&audioID=573924*
11 *Netzpolitik.org v. 14.4., 27.4 und 19.6.2015, https://netzpolitik.org/2015/bka-datenbank-bundesdatenschutzbeauftragter-fand-gravierenden-verstoss-gegen-datenschutzrechtliche-vorschriften/, https://netzpolitik.org/2015/innenministerium-bestaetigt-rechtswidrige-speicherung-linker-aktivistinnen/, https://netzpolitik.org/2015/nachhilfe-der-bundesdatenschutzbeauftragten-fuehrt-zu-90-schwund-in-polizeidatenbank-zu-linkem-aktivismus/*
12 *Beispielsweise § 98a StPO, § 47 ASOG (Berlin) oder § 28 BKAG*
13 *Tagesschau v. 3.10.2017, https://www.tagesschau.de/inland/g20-akkreditierungen-107.html*
14 *Tagesschau v. 30.08.2017, https://www.tagesschau.de/inland/gzwanzig-datenschuetzer-101.html*
15 *Tagesspiegel v. 1.9.2017, https://www.tagesspiegel.de/politik/kriminalitaetsdatenbanken-bka-praesident-wehrt-sich-gegen-speichervorwuerfe/20273314.html*
16 *Tagesschau v. 3.10.2017, https://www.tagesschau.de/inland/g20-akkreditierungen-107.html*
17 *Zeit.de v. 30.8.2017, http://www.zeit.de/gesellschaft/zeitgeschehen/2017-08/datenschutz-datenspeicherung-bka-heiko-maas-rechtswidrig-aufklaerung*
18 *Deutschlandradio, Informationen am Morgen v. 2.9.2017, http://srv.deutschlandradio.de/themes/dradio/script/aod/index.html?audioMode=3&audioID=573924*
19 *Guardian v. 10.10.2013, https://www.theguardian.com/commentisfree/2013/oct/10/double-danger-nsa-surveillance*
20 *Beschreibung der datengetriebenen Hoffnung in der Politik: Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung, Pressemitteilung v. 21.12.2016, https://www.fiff.de/presse/pressemitteilungen/digitalcharta-notwendige-politische-initiative-trotz-grober-fehler-fiff-sichert-mithilfe-zu*
21 *Deutschlandfunk-Kultur, Weltzeit v. 5.9.2017, http://www.deutschlandfunkkultur.de/chinas-sozialkredit-system-auf-dem-weg-in-die-it-diktatur.979.de.html?dram%3Aarticle_id=395126*
22 *zeit.de v. 15.10.2015, http://www.zeit.de/politik/2015-10/nsa-affaere-untersuchungsausschuss-metadaten-brandon-bryant-aussage/komplettansicht*
23 *zeit.de v. 16.10.2015, http://www.zeit.de/politik/ausland/2015-10/usa-drohnen-drohnenkrieg-rechtfertigung/komplettansicht*
24 *Netzpolitik.org v. 3.7.2014, https://netzpolitik.org/2014/nsa-ueberwacht-tor-infrastruktur-und-alle-nutzer-auch-betreiber-in-deutschland/ und Theintercept.com v. 18.2.2014, https://theintercept.com/2014/02/18/snowden-docs-reveal-covert-surveillance-and-pressure-tactics-aimed-at-wikileaks-and-its-supporters/*
25 *Siehe dazu auch die rassistische Rückfall-Vorhersagesoftware https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing*
26 *Netzpolitik.org v. 6.10.2017, https://netzpolitik.org/2017/pre-crime-ueber-menschen-die-ungewollt-teil-von-datenexperimenten-sind/*
27 *Netzpolitik.org v. 23.5.2017, https://netzpolitik.org/2017/funkzellenabfrage-letztes-jahr-landeten-handy-daten-aller-berliner-alle-elf-tage-bei-der-polizei/*
28 *sueddeutsche.de v. 6.7.2015, http://www.sueddeutsche.de/wirtschaft/angst-vor-ueberwachung-die-hilfssheriffs-der-telekom-1.2551588*
29 *heise.de v. 5.1.2017, https://www.heise.de/newsticker/meldung/Googles-KI-AlphaGo-gewinnt-und-gewinnt-3589295.html*

Markus Reinisch

Vermessen, berechnen und vorhersagen

Zahlengläubigkeit und positivistisches Grundverständnis von *Big Data*

Algorithmen *und* Deep Learning, kognitive Roboter, intelligente *Maschinen, vernetzte* Smart *Things (Internet der Dinge) sind nur einige der derzeit diskutierten Schlagworte, wenn es um die technische, insbesondere digitale Beschleunigung geht. Und es kommen laufend neue Schlagworte hinzu. Sie gehen einher mit der kontroversen Big-Data-Debatte, die nicht mehr nur in den Fachwissenschaften und Feuilletons geführt wird. Während Technologie-Euphoriker in Big Data und der zunehmenden* Datafizierung *von Lebensbereichen eine revolutionäre Art des Erkenntnisgewinns, von Vorhersagemethode und Effizienzsteigerung sehen, werden immer mehr kritische Stimmen laut. Vor allem solche, die sich gegen die fortschreitende Ausrichtung am Vermessen und damit am Quantitativen richten. Es wird heute in vielen Lebensbereichen eine unüberschaubare Menge an Daten erfasst, analysiert und tabellarisiert, um passgenaue Vorhersagen zu erreichen, beispielsweise für menschliches Verhalten. Hinter der Zahlengläubigkeit steckt jedoch eine positivistische Weltsicht, die alles auszuklammern scheint, was nicht mess- und formalisierbar ist. Dass dies nicht ohne gesellschaftliche, politische, ethische und bildungstheoretische Folgen bleiben wird, soll dieser Beitrag zeigen.*