ATHENE
Nationales Forschungszentrum
für angewandte Cybersicherheit

Systematic Privacy for large, real-life Data Processing Systems

AUTOREN

Amina Gutjahr
Goethe Universität Frankfurt

Till Schaller
Universität Kassel

Gerrit Hornung
Universität Kassel

Annika Selzer
Fraunhofer SIT

Sarah Stummer
Fraunhofer SIT

Jessica Kriegel
Fraunhofer SIT

Indra Spiecker gen. Döhmann
Goethe Universität Frankfurt

Thomas Wilmer
Hochschule Darmstadt

Systematic Privacy for large, real-life Data Processing Systems

Impressum

Kontakt
Nationales Forschungszentrum für angewandte Cybersicherheit ATHENE
c/o Fraunhofer-Institut für Sichere Informationstechnologie SIT
Rheinstraße 75
64295, Darmstadt

Hinweise

Dieser Beitrag wurde mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) im Rahmen ihrer gemeinsamen Förderung für das Nationale Forschungszentrum für angewandte Cybersicherheit ATHENE unterstützt.

Autoren

Amina Gutjahr
Goethe Universität Frankfurt

Gerrit Hornung
Universität Kassel

Jessica Kriegel
Fraunhofer SIT

Till Schaller
Universität Kassel

Annika Selzer
Fraunhofer SIT

Indra Spiecker gen. Döhmann
Goethe Universität Frankfurt

Sarah Stummer
Fraunhofer SIT

Thomas Wilmer
Hochschule Darmstadt

Inhalt

Vorwort

Die vorliegende Studie ist im Rahmen des Projekts „Systematic Privacy for large, real-life Data Processing Systems“ entstanden. Das Projekt wird im Forschungsbereich Legal Aspects of Privacy and IT Security (LeAP) bei ATHENE, dem Nationalen Forschungszentrum für angewandte Cybersicherheit und Europas größtem Forschungszentrum für Cybersicherheit und Privatsphärenschutz, durchgeführt. An dem Projekt beteiligt sind vier im Bereich des Digitalisierungsrechts führende hessische Lehrstühle/Forschungseinrichtungen: Prof. Dr. Gerrit Hornung, LL.M. (Edinburgh) und Till Schaller, Universität Kassel (UKS), Dr. Annika Selzer, Sandra Stummer, LL.M. und Jessica Kriegel, Fraunhofer SIT, Forschungsabteilung „IT Law & Interdisciplinary Privacy Research“ (SIT), Prof. Dr. Indra Spiecker genannt Döhmann, LL.M. (Georgetown University), und Amina Gutjahr, Goethe Universität Frankfurt (GUF) und Prof. Dr. Thomas Wilmer, Hochschule Darmstadt (h_da).

Im Projekt befassen wir uns mit den besonderen Herausforderungen, die große Daten–sätze, insbesondere solche aus verschiedenen Quellen und in verschiedenen Formaten, an das Recht stellen. Im besonderen Fokus stehen Datenschutz und IT-Sicherheit, aber auch das Urheber- und Geheimnisschutzrecht wird bearbeitet. Naturgemäß können aus der Vielzahl der Problemlagen nur einzelne Teilbereiche herausgegriffen werden, um hierfür schlüssige Konzepte für die Anwendung zu entwickeln. Damit wollen wir einen Beitrag dazu leisten, dass eine rechtssichere datenschutz- und IT-sicherheitsrechtskonforme Nutzung der Verwertung von großen Datenbeständen möglich wird.

Die vorliegende Untersuchung stellt aus den vier Teilprojekten wesentliche Zwischenergebnisse in Zusammenhang. Wir befassen uns mit den besonderen Herausforderungen beim Einsatz von Verhaltensbeeinflussung – wie sie erst durch die systematische Verwendung von großen Datenmengen möglich wird – durch den Staat, klären ab, welche urheberrechtlichen Probleme entstehen können und wie die KI-Verordnung auf große Datensätze reagiert, ob und wie Anonymität ein Instrument sein kann, um Datenschutzkonformität in sich ständig verändernden großen Datensätzen herzustellen, und wie die Schnittstelle von IT-Sicherheit und Datenschutz mithilfe technischer und organisatorischer Maßnahmen konkretisiert werden kann.

Darmstadt, Frankfurt und Kassel, im Oktober 2023

1. Einleitung

In der Digitalisierung bestimmen vernetzte Computersysteme zunehmend unser tägliches Leben. Die mit dieser Technisierung einhergehende Verarbeitung von Daten ist ubiquitär: Überall werden Daten erhoben, gespeichert und ausgewertet. Aufgrund der Digitalisierung nahezu aller Lebensbereiche, von der Alltags-Kommunikation über das öffentliche und private Verkehrswesen, „Smart Cities" und die Arbeitswelt bis hin zur öffentlichen Verwaltung und dem Gesundheitswesen, ist nicht mehr nur vom „Internet of Things", sondern vom „Internet of Everything" die Rede.[1]

Die Grundlage eines jeden Computersystems zur Auswertung und Nutzung dieser Daten bilden Algorithmen. Ein Algorithmus ist eine Folge wohldefinierter Rechenschritte, die eine Eingabe in eine Ausgabe umwandeln.[2] Sie werden unter anderem – und zunehmend – dazu eingesetzt, aus großen, komplexen Datensätzen unterschiedlicher Herkunft, Sortierung und Qualität („Big Data") Muster, Korrelationen und Wahrscheinlichkeiten zu erkennen und neue Erkenntnisse zu gewinnen. Big Data zeichnet sich dadurch aus, dass eine große Menge („Volume") qualitativ sehr unterschiedlicher Arten von Daten, die aus vielfältigen Quellen erhoben und gespeichert wurden („Variety"), in hoher Geschwindigkeit („Velocity") verarbeitet werden, um daraus neue Informationen zu ziehen und diese wertschöpfend fruchtbar zu machen („Value").[3] Weitere Charakteristika hat die European Union Agency For Network And Information Security (ENISA) 2015 herausgearbeitet:[4] verteilte und redundante Datenspeicherung, parallele Aufgabenbearbeitung, Skalierbarkeit, Hardwareunabhängigkeit, geringe Kosten. Sowohl für die Wirtschaft als auch für den Staat gilt, dass die Anhäufung großer Datenvolumina inhaltlicher Varietät, die in höchster Geschwindigkeit – mit teils sich selbst fortentwickelnden Algorithmen – ausgewertet werden, eine nie dagewesen Chance auf Informationsgewinnung bietet.[5] Häufig werden die Daten dabei nicht mehr zu dem Zweck verwendet, zu dem sie ursprünglich erhoben wurden.[6] Allerdings handelt es sich bei den verarbeiteten Daten nicht unbedingt um personenbezogene Daten; je nach Verarbeitungskontext und -ziel können anonymisierte Daten für die Big-Data-Analyse ausreichen.[7]

Die systemische Analyse von massenhaften Daten, von Big Data, kann auf unterschiedliche Weise erfolgen. Vereinfacht lässt sich das so darstellen: Bei der sog. deskriptiven Analytik werden Datensätze vergangenheitsbezogen ausgewertet, um daraus empirische Erkenntnisse zu ziehen. Diese Vorgehensweise kommt z.B. – jedoch nicht ausschließlich – beim Data Mining[8] zum Einsatz. Die auf diese Weise erkannten (Verhaltens-) Muster und

[1] Vgl. *Hofstetter*: Das Ende der Demokratie, S. 28.
[2] *Cormen/Leiserson/Rivest/Stein/Molitor:* Algorithmen, S. 5; *Russell/Norvig*: Artificial Intelligence, S. 25 ff.
[3] *Laney*, „3D Data Management: Controlling Data Volume, Velocity, and Variety", Gartner, File Nr. 949, https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf; ähnliche Merkmale nennt auch die *Länderarbeitsgruppe „Digitaler Neustart"*, Bericht zu Big Data, S. 8; auch der *Wissenschaftliche Dienste des Deutschen Bundestags*, Aktueller Begriff – Big Data, https://www.bundestag.de/resource/blob/194790/c44371b1c740987a7f6fa74c06f518c8/big_data-data.pdf; *Schulz* in *Gola/Heckmann (Hrsg.)*, DSGVO/BDSG, Art. 6 Rdnr. 151; *Hoffmann-Riem* in *ders. (Hrsg.):* Big Data, S. 19 f.
[4] *ENISA,* Big Data Security, S. 8 f.
[5] *DSK*, Entschließung: Big Data zur Gefahrenabwehr und Strafverfolgung: Risiken und Nebenwirkungen beachten, https://www.datenschutzkonferenz-online.de/media/en/20150318_en_BigData.pdf.
[6] Die Rechtmäßigkeit richtet sich daher meist nach Art. 6 IV DSGVO, *Roßnagel* in *Simitis/Hornung/Spiecker gen. Döhmann (Hrsg.)*, DSGVO/BDSG, Art. 6 DSGVO, Rdnr. 12 ff.
[7] *Schantz* in *Simitis/Hornung/Spiecker gen. Döhmann (Hrsg.)*, DSGVO/BDSG, Art. 6 Rdnr. 122; *Schulz* in *Gola/Heckmann (Hrsg.)*, DSGVO/BDSG, Art. 6 Rdnr. 152 f; European Commission, Questions and Answers – Data protection reform.
[8] Data Mining beschreibt ein systematisches Verfahren, bei dem mittels statistischer Methoden aus großen Datenbeständen Querverbindungen und Zusammenhänge der Daten hergestellt und anhand dessen bestimmte Muster oder Trends erkannt werden sollen. Es sollte bei der Antiterrordatei zum Einsatz kommen (die entsprechende Vorschrift des § 6a II 2 1 ATDG wurde aber durch BVerfGE 156, 11 – „Antiterrordatei II" für nichtig erklärt), zur Definition von Data Mining vgl. die Antwort der Bundesregierung auf eine Anfrage der Abgeordneten Jelpke u.a. und der Fraktion DIE LINKE, *Deutscher* Bundestag, BT-Drs. 17/11582, S. 3.

Querverbindungen können sodann im Wege der prädiktiven Analytik dazu genutzt werden, künftiges Handeln zu prognostizieren. So kann beispielsweise errechnet werden, wie wahrscheinlich es ist, dass ein Kunde[9] sich durch eine bestimmte Werbung dazu verleiten lassen wird, ein bestimmtes Produkt zu kaufen, ein Schüler einen Schulabschluss schafft, ein Bürger eine bestimmte Partei wählt, oder welche Chancen eine bestimmte Person auf dem Arbeitsmarkt hat.[10] Die sog. präskriptive Analytik verknüpft deskriptive und prädiktive Analytik und ermöglicht durch die Kombination abbildender Erkenntnisse und Prognosen, Handlungsempfehlungen zu geben und das erlangte Wissen strategisch zur Erreichung bestimmter Ziele einzusetzen.[11]

Außerdem können Daten mithilfe dieser Analysemethoden dazu genutzt werden, Persönlichkeitsprofile zu erstellen. Diese ermöglichen es, anhand der Analyse von Interessen, Erfahrungen, sozialer Einbindung, Voreinstellungen und Verhaltensweisen künftiges Verhalten von Individuen in vielfältigster Weise zu prognostizieren. So kann die Persönlichkeit eines Individuums umfassend psychometrisch vermessen und Personen in bestimmte Gruppen kategorisiert werden. Da es sich bei den in großen Datenverarbeitungssystemen verarbeiteten Daten oft um personenbezogene Daten handelt, birgt deren Verarbeitung, insbesondere aufgrund des technologischen Fortschritts und der zunehmenden Konnektivität, Risiken für die Rechte und Freiheiten der betroffenen Personen. Insbesondere werden einzelne Personen, ihr Verhalten und ihre Interessen sowie ihre Verbindung zu anderen Personen nachvollziehbar. Werden Personen bei der Zuordnung zu einem bestimmten Persönlichkeitsprofil z.B. Eigenschaften, Verhaltensweisen oder Interessen zugeschrieben, wird – unabhängig davon, ob sie sich selbst mit den Zuschreibungen identifizieren oder nicht – einer Vielzahl von Diskriminierungen der Weg geebnet.[12] Außerdem können die in großen Datenverarbeitungssystemen ausgewerteten personenbezogenen Daten zum Identitätsdiebstahl genutzt werden und dadurch zu finanziellen oder gesellschaftlichen Nachteilen[13] führen.

Die umfassenden Möglichkeiten zur Verknüpfung und Generierung neuer Erkenntnisse und Prognosen aus großen Datensätzen machen deutlich, dass solche Datensätze überall dort, wo sie erhoben, gespeichert und ausgewertet werden, grundsätzlich auch dafür verwendet werden können, das Verhalten von Personen zu beeinflussen. Zwar hat auch das Recht selbst bereits eine verhaltenssteuernde Wirkung und wird auch gezielt dazu eingesetzt. Werden zur Generierung solcher rechtlichen Maßstäbe jedoch personenbezogene Daten von Bürgern ausgewertet und daraus Rückschlüsse gezogen, derer sich der Einzelne gar nicht bewusst ist (und oft genug nicht sein kann), entsteht ein ganz neues Gefährdungspotenzial. Und selbst die Kenntnis der Auswertung genügt oftmals nicht, um ein darüber entstehendes Machtungleichgewicht aufzuheben.[14]

Während Big-Data-Anwendungen sich im kommerziellen Bereich bereits zu einer effektiven und weit verbreiteten Methode zur Analyse und Beeinflussung bis hin zur Manipulation[15] des Konsumverhaltens entwickelt haben, erst recht im Online-Marketing[16], sind sie im öffentlichen Sektor – zumindest in Deutschland – bislang weniger umfassend etabliert.

---

[9] Aus Gründen der besseren Lesbarkeit wird in der vorliegenden Studie auf die Verwendung männlicher und weiblicher Sprachformen verzichtet. Es wird das generische Maskulinum verwendet, wobei alle Geschlechter gleichermaßen gemeint sind.

[10] Zur prädiktiven Analytik *Mühlhoff*: Predictive Privacy; zum Arbeitsmarktservice Österreich (AMS), der arbeitslose Personen nach ihren Chancen auf dem Arbeitsmarkt kategorisiert, *Fröhlich/Spiecker gen. Döhmann*, Können Algorithmen diskriminieren?, https://verfassungsblog.de/koennen-algorithmen-diskriminieren/.

[11] Zum Vorstehenden *Hoffmann-Riem*, AöR 2017, 1 (7 f.).

[12] *Britz*: Einzelfallgerechtigkeit versus Generalisierung, S. 75 ff.; zu Diskriminierung durch Algorithmen außerdem *Martini*, JZ 2017, 1017 (1018); *Härtel*, LKV 2019, 49 (56).

[13] *DSK*, Kurzpapier Nr.18, S.3; *Heberlein* in *Ehmann/Selmayr*, DSGVO, Art. 6 Rdnr. 28; *Martini* in *Paal/Pauly*, DSGVO/BDSG, Art. 24 Rdnr. 29a.

[14] Allgemein zur Verhaltenssteuerung durch Algorithmen *Hoffmann-Riem* in *ders. (Hrsg.)*:Rechtliche Rahmenbedingungen für und regulative Herausforderungen durch Big Data, S. 23 f.

[15] Z.B. im Falle von Dark Patterns, s. dazu unten, S.14.

[16] Dazu *Borah/Skiera*, IJRM 2021, Vol. 38, Issue 4, S. 811-816; vgl. *Hoffmann-Riem,* in *ders. (Hrsg.)*: Big Data, S. 23 f.

In anderen Ländern sieht das aber schon ganz anders aus. Im Nachbarland Österreich z.B. wurde schon 2019 eine Software eingesetzt, welche arbeitslose Personen nach ihren „Arbeitsmarktintegrationschancen", das heißt ihren Perspektiven auf einen Arbeitsplatz, kategorisiert und an diese Kategorisierung Fördermaßnahmen knüpft. Der Einsatz sorgte für große Furore.[17] Mit neueren Entwicklungen wie der elektronischen Patientenakte (ePA) in Verbindung mit dem im Digitale Versorgunggesetz (DVG) vorgesehenen Forschungshub, halten Big-Data-Auswertungen aber dennoch auch in Deutschland schleichend Einzug in die Gerichts- und Verwaltungspraxis und damit in Bereiche unmittelbaren staatlichen Handelns.

Neben den unklaren Grenzen staatlicher Verhaltensbeeinflussungen auf Grundlage der Auswertung von Big Data bringt die Verarbeitung großer Datenmengen auch eine ganze Reihe von Fragen aus dem Bereich des geistigen Eigentums und gewerblicher Schutzrechte mit sich. Dies betrifft sowohl die Datengewinnung, welche geschützte Werke und Datenbanken umfassen kann, als auch die Schutzfähigkeit der Ergebnisse der Datennutzung. Gesetzlich vorgesehen sind besondere Erlaubnisse zur Datennutzung im Text- und Data-Mining-Bereich sowie in der Forschung nach den §§ 44b, 60d UrhG, die jedoch nur für bestimmte Anwendungsbereiche in Betracht kommen[18]. Neben der Absicherung der Datensätze vor entsprechenden urheberrechtlichen Unterlassungsansprüchen können sich auch haftungsrechtliche Herausforderungen stellen. Hier können fehlerhafte Datensätze zu Folgefehlern in der KI-Anwendung führen. Je nach Einsatzzweck und möglicher umfassender Integration der KI in Netzwerke ist schwer abschätzbar, welche Schäden typischerweise zu erwarten sein können[19]. Vor allem beim Einsatz großer Datenmengen in Hochrisiko-KI-Systemen sind die regulatorischen Vorgaben der künftigen KI-Verordnung und der KI-Haftungsverordnung zu beachten. Neue Offenlegungspflichten werden im Hinblick auf die Opazität und Autonomie der KI gerade bei großen Datenmengen schwer umzusetzen sein und Probleme bei der Ermittlung der Entscheidungs- und Datenlieferkette aufwerfen.

Angesichts der weitreichenden Risiken für die Bürger gibt es zahlreiche Initiativen zum Schutz vor den manipulativen Aspekten der Verarbeitung großer Datensätze. Allerdings ist zu konstatieren, dass – mit Ausnahme des KI-VO-Entwurfs der EU – in der Regel nicht gezielt auf die besonderen Gefahren der Auswertung großer Datensätze abgestellt wird, sondern ganz generell Beschränkungen vorgesehen werden. So verfolgt auch die DSGVO weitgehend – Ausnahmen bestehen insbesondere in Bezug auf die Umsetzung technischer und organisatorischer Maßnahmen[20] – keinen risikobasierten Ansatz,[21] der womöglich danach unterscheidet, in welchem Kontext und in welcher Zusammenführung Daten verwendet werden. Dies ist im Hinblick auf die Datenverarbeitung als dem ersten Schritt einer Datenauswertungskette massenhafter Datensätze in algorithmischen Systemen auch ein zulässig selbstbeschränkender und angesichts der vielfachen Möglichkeiten des Einsatzes von Daten[22] kluger Ansatz dieses Rechtsregimes. Er reicht aber nicht, um die Gefahrenlagen, die für Einzelne entstehen, angemessen aufzufangen und unerwünschte Machtasymmetrien zu vermeiden. Vielmehr muss dafür auch eine Regulierung der Zwischenschritte der Auswertung und Bearbeitung bis hin zur Anwendung dieser Ergebnisse greifen, um die Gefahrenpotenziale einzuhegen.[23]

[17] Dazu *Fröhlich/Spiecker gen. Döhmann,* Können Algorithmen diskriminieren?, https://verfassungsblog.de/koennen-algorithmen-diskriminieren/.
[18] *Jacobsen/Hartmann,* MMR-Aktuell 2021, 441332.
[19] Siehe zu den Risiken *Zech,* ZfPW 2019, S. 198ff.
[20] S. dazu unten, S. 35.
[21] Der Vorschlag eines risikobasierten Ansatzes hat sich in der DSGVO nicht durchgesetzt, zumindest nicht in Bezug auf die generelle materielle Rechtmäßigkeit der Verarbeitung. Dennoch sind aber einige Anforderungen in der DSGVO von einer Risikobewertung abhängig (so z.B. Art. 24, 25 und 32 DSGVO), *Hornung/Spiecker gen. Döhmann* in: Simits/ders./dies. (Hrsg.), DSGVO/BDSG, Einleitung, Rn. 242 f.
[22] *Spiecker gen. Döhmann* in *Kischel/Kube (Hrsg.),* HStR[4], § 20 Rdnr. 26 ff.
[23] *Spiecker gen. Döhmann* in *Kischel/Kube (Hrsg.),* HStR[4], § 20, Rdnr. 54 f.

Die Ansätze von einzelnen Gegenmaßnahmen, die über systematische Beschränkungen und Rahmensetzungen hinausgreifen können, assistieren weiter. Um die Risiken für die Rechte und Freiheiten der betroffenen Personen zu minimieren, könnten personenbezogene Daten, wann immer möglich, anonymisiert werden. Aus rechtlicher Sicht sind die Anforderungen an die Anonymisierung in der DSVGO recht vage gehalten. So unternimmt lediglich Erwgr. 26 DSGVO einen Versuch der Begriffsdefinition als „Informationen, die sich nicht [oder nicht mehr] auf eine identifizierbare natürliche Person beziehen". Infolgedessen werden in der Datenschutzpraxis sowie in der juristischen Fachliteratur sowohl die Begriffe „anonym" und „Anonymisierung" als auch die rechtlichen Anforderungen an die Anonymisierung umfassend analysiert und diskutiert,[24] insbesondere im Hinblick darauf, ob der Begriff „anonym" absolut oder relativ zu verstehen ist.[25] Diese Fachdiskussionen führen i.d.R. zu dem Schluss, dass die Anforderungen an die Anonymisierung nach der DSGVO nicht spezifisch genug sind, um die Anonymisierung in der Praxis rechtssicher anzuwenden.[26] Gründe hierfür werden u.a. in der Ausweitung des Begriffs „personenbezogene Daten" auf Daten, die zunächst nicht personenbezogen erscheinen,[27] fehlenden Prozessen zur Anonymisierung und mangelnden Möglichkeiten zur Bewertung der Anonymität von Daten gesehen.[28] Darüber hinaus werden immer wieder unzureichend umgesetzte Anonymisierungen sowie De-Anonymisierungspotenziale aufgedeckt.[29]

Soweit Daten personenbezogen sind, greifen auch datenschutzrechtliche Vorgaben für ihre technische Sicherung. Big Data wirft aber unabhängig vom Personenbezug Fragen der rechtlichen Regulierung der IT-Sicherheit auf. Auch wenn sich also Umgebungen und Akteure in Big-Data-Anwendungen unterscheiden, so wird die facettenreiche Thematik in der Frage nach Sicherheit (sog. „Big Data Security") dennoch vereint,[30] denn Datenkonglomerate stellen ein attraktives Ziel für Angreifer dar. Die Erlangung von Daten zu Spionage- oder auch Erpressungszwecken gehört zu den wesentlichen Zielen der Täter, [31] die in großen Datenansammlungen ein noch attraktiveres Ziel finden. Besonders bedroht sind dabei personenbezogene Daten, die im Falle einer Beeinträchtigung der Sicherheit eine erhebliche Gefährdung oder sogar Schädigung für die Grundrechte natürlicher Personen bedeuten können.[32] Allein der Blick auf die Verarbeitungsprozesse, in denen personenbezogene Daten vorkommen, würde aber den Gefahren von Cyberangriffen nicht gerecht werden, die durch Big-Data-Anwendungen noch potenziert werden, da die Angriffe nicht immer unmittelbar in der Datenverarbeitung ansetzen und häufig über verschiedene Kanäle von innen und außen erfolgen.

Ein lückenloser, das Datenschutz- und IT-Sicherheitsrecht übergreifender Rechtsrahmen, der die Besonderheiten von Big-Data-Anwendungen berücksichtigt und auf den Schutz personenbezogener Daten ausgerichtet ist, fehlt bislang.[33] Angesichts der sich teilweise überschneidenden Anwendungsbereiche der beiden Rechtsgebiete und ihrer Ausrichtung auf unterschiedliche Schutzgegenstände ist dies – auch in Anbetracht der vielschichtigen

[24] *European Data Protection Board*, *Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak*, S.8 ff.; *Article 29 Data Protection Working Party*, *Opinion 4/2007 on the concept of personal data*, S.9; *Data Protection Commission*, *Guidance on Anonymisation and Pseudonymisation*, https://www.dataprotection.ie/sites/default/files/uploads/2020-09/190614%20Anonymisation%20and%20Pseudonymisation.pdf, S.14.
[25] *Bischoff*, PharmR 2020, 309 (313); *Gierschmann*, ZD 2021, 482 (483); *Marnau,* DuD 2016, 428 (429f.).
[26] *Winter/Battis/Halvani*, ZD 2019, 489 (493); *Marnau/Berrang/Humbert*, DuD 2018, 83 (88); *Selzer/Timm*, DuD 2021, 816 (816).
[27] *Ohm*, UCLA Law Review 2010, 1701 (1701 ff.); vgl. *Roßnagel*, ZD 2018, 243 (244); *Herbst*, NVwZ 2016, 902 (903 f.).
[28] *Selzer/Timm*, DuD 2021, 816 (816); vgl. *Winter/Battis/Halvani*, ZD 2019, 489 (490).
[29] *Narayanan/Shmatikov*, IEEE Symposium on Security and Privacy 2008, 111 (112); *De Montjoye/Hidalgo/Verleysen/Blondel*, SCIENTIFIC REPORT 2013, 1 (1f.); *Bohannon*, SCIENCE 2013, 262 (262).
[30] *ENISA,* Big Data Security, „One of the main issues in using Big Data systems is security", S. 4.
[31] *BSI,* BSI-Lagebericht 2022.
[32] *Hornung* in *Hoffmann-Riehm*: Big Data – Regulative Herausforderungen, S. 81 ff.; *Hornung/Herfurth* in König/Schröder/Wiegand, Big Data – Chancen, Risiken, Entwicklungstendenzen, 149 ff.; *Roßnagel/Nebel,* DuD 2015, 455.
[33] *Schulz* in Gola/Heckmann, DSGVO/BDSG, Art. 6 Rdnr. 152.

Umgebungen und Akteure – überdies höchst voraussetzungsvoll. Dazu trägt maßgeblich auch die starke Fragmentierung der IT-Sicherheitsbestimmungen in den verschiedenen Regelungen des Datenschutzrechts sowie des allgemeinen und sektorspezifischen IT-Sicherheitsrechts bei.

Für den öffentlich-rechtlichen und vom Normumfang bedeutendsten Teil des IT-Sicherheitsrechts muss konstatiert werden, dass dieser zwar wichtige, zahlenmäßig aber bislang nur wenige Unternehmen zu IT-Sicherheitsmaßnahmen verpflichtet. Neben Betreibern kritischer Infrastrukturen werden Anbieter digitaler Dienste und – seit dem IT-Sicherheitsgesetz 2.0[34] – auch Unternehmen im besonderen öffentlichen Interesse reglementiert. Unterhalb der Schwelle regulierter Unternehmen, wozu insbesondere KMU zählen, müssen die IT-Sicherheitsmaßnamen meist nur fakultativ umgesetzt werden, was zu unzureichenden Schutzmaßnahmen führt.[35] Im Übrigen umfasst die Querschnittmaterie des IT-Sicherheitsrechts auch zivilrechtliche Bestimmungen. Jedoch machen weder die Organisationspflichten des Gesellschaftsrechts noch die allgemeinen Haftungsbestimmungen, die zunächst auf alle Kapitalgesellschaften Anwendung finden, konkrete Vorgaben für ein Mindestmaß an IT-Sicherheit.[36]

Im Datenschutzrecht, dessen Anwendungsbereich die Verarbeitung personenbezogener Daten erfasst, besteht zwar mit Art. 32 DSGVO eine generalklauselartige Pflicht für sämtliche Organisationen, angemessene technisch-organisatorische Maßnahmen zum Schutz der personenbezogenen Daten zu implementieren; dies wird durch die allgemeinen Regelungen zu Verantwortung und Rechenschaft in Art. 24 und Art. 5 Abs. 2 DSGVO ergänzt. Jedoch fehlt es bislang an einer hinreichenden Operationalisierung des inhärenten risikobasierten Ansatzes in Art. 32 DSGVO für die Anforderungen der Big Data; dies gilt vor allem unter Berücksichtigung der zugrundeliegenden IT-Systeme solcher Unternehmen, wie viele KMU, die bislang keiner IT-Sicherheitsregulierung unterfallen.

Daher bedarf es für den Schutz personenbezogener Daten in Big-Data-Anwendungen eines holistischen Ansatzes, der auch die zugrundeliegende IT-Infrastruktur erfasst und ein harmonisiertes Rechtsgerüst zwischen Datenschutz und IT-Sicherheit schafft. Ein Schutz personenbezogener Daten ist ohne IT-Sicherheit nicht mehr zu gewährleisten.[37] Dem Recht wird dabei die Aufgabe zuteil, die Anforderungen einheitlich und replizierbar zu gestalten. Angesichts dieser verschiedenen Problemkreise im Zusammenhang mit der Verarbeitung großer Datensätze setzt sich das erste Teilkapitel der vorliegenden Studie damit auseinander, inwiefern und in welchen Bereichen des öffentlichen Sektors die Ergebnisse von Big-Data-Analysen überhaupt zur Verhaltensbeeinflussung genutzt werden können (2.). Sodann wird das Problem der Rechtsunsicherheiten im Hinblick auf Schutzrechte an Daten behandelt und inwiefern kommende EU-Regelungen wie der KI-VO-E, die KI-Haftungsrichtlinie sowie die Produkthaftungslinie mehr Klarheit in Bezug auf das Eigentums- und Urheberrecht an Daten schaffen könnten (3.). Das nächste Teilkapitel befasst sich mit den rechtlichen Anforderungen an die Anonymisierung personenbezogener Daten als einer Strategie, zumindest die Personenbezogenheit des Auswertungsmaterials aufzulösen (4.). Das letzte Teilkapitel setzt sich mit den Gefahren sowie dem bestehenden Rechtsrahmen für die IT-Sicherheit in Big-Data-Systemen auseinander und zeigt, wie das bestehende Schutzniveau optimiert werden könnte (5.). Abschließend werden die wesentlichen Ergebnisse der Studie zusammengefasst (6.).

[34] Zweites Gesetz zur Erhöhung der Sicherheit informationstechnischer Systeme, BGBl. 2021 I, S. 1122; näher *Hornung,* NJW 2021, 1985; *Schallbruch*, CR 2021, 450; *ders*., CR 2021, 516.
[35] *Hillebrand/Niederprüm/Schäfer/Thiele/Henseler-Unger,* Aktuelle Lage der IT-Sicherheit in KMU, S. 88 f.
[36] *Spindler,* Verantwortlichkeiten von IT-Herstellern, Nutzern und Intermediären, S. 297, konstatiert: „einheitliche Regelungen oder ein übergreifender Ansatz zur Gewährleistung einer grundlegenden IT-Sicherheit – unabhängig von vertraglichen Regelungen – fehlen".
[37] *Hornung/Schallbruch* in *Hornung/Schallbruch*: IT-Sicherheitsrecht, § 1 Rdnr. 17.

2. Verhaltensbeeinflussung durch Big-Data-Analysen im öffentlichen Sektor

Die Grundlage eines jeden Computersystems bilden Algorithmen. Algorithmen sind Rechenoperationen, die in einer determinierten Abfolge durchgeführt werden, um mithilfe eingespeister Daten genau definierte Probleme zu lösen.[38] Übertragen in binären Maschinencode sind sie die Basis einer jeden Software und ermöglichen die selbsttätige Umsetzung des Programmcodes, eingespeiste Informationen (Input) in einen Output zu verwandeln.[39] Wie auch das Recht sind Algorithmen geeignet und werden häufig gezielt dafür eingesetzt, menschliches Verhalten zu steuern und zu beeinflussen und die gesellschaftliche Ordnung zu gestalten.[40] Sie werden dazu genutzt, unsere Interessen und unser Verhalten zu analysieren. Anhand dieser Daten kann sodann unser künftiges Handeln prognostiziert werden. Um eine möglichst präzise Prognose zu generieren, werden immer größere Datensätze gesammelt und ausgewertet.

Gerade wenn sich der Staat Big-Data-Analysen zur direkten Verhaltensbeeinflussung bedient, muss aber besonderes Augenmerk darauf gelegt werden, nicht über Gebühr in die Freiheitssphäre der Bürger einzugreifen. Andererseits darf auch nicht vergessen werden, dass auch dem Recht eine verhaltenssteuernde Wirkung zukommt, die gerade auch in der „neuen Verwaltungsrechtswissenschaft"[41] analysiert wird. Daher gilt es, die spezifischen Risiken einer algorithmengesteuerten Beeinflussung herauszuarbeiten. Im Folgenden wird dazu zunächst dargestellt, auf welche Art und Weise eine Steuerung und Beeinflussung menschlichen Verhaltens durch die Ergebnisse von Big-Data-Analysen prinzipiell möglich ist (2.1) und inwiefern solche Beeinflussungen Grundrechtsrelevanz aufweisen (2.2). Sodann werden Szenarien aufgezeigt, in denen die Ergebnisse von Big-Data-Analysen in verschiedenen Bereichen des öffentlichen Sektors insbesondere im internationalen Vergleich bereits konkret zur Verhaltensbeeinflussung genutzt werden oder genutzt werden könnten (2.3), bevor in einem abschließenden Fazit die Befunde ausgewertet werden (2.4).

2.1 Formen der Verhaltensbeeinflussung durch Big-Data-Analysen

Es gibt verschiedene Formen, wie die Ergebnisse von Big-Data-Analysen genutzt werden könnten, um individuelles Verhalten zu steuern und zu beeinflussen.

2.1.1 Entscheidungsarchitektur

Zunächst einmal können die Ergebnisse der Big-Data-Analyse zur Gestaltung einer Entscheidungsarchitektur genutzt werden. Die US-amerikanischen Wissenschaftler *Richard*

[38] *Knorre*, Big Data im öffentlichen Diskurs, in: Knorre/Müller-Peters/Wagner, Die Big-Data-Debatte, 2020, S. 1 (6); *Spiecker gen. Döhmann/Towfigh*, Das Allgemeine Gleichbehandlungsgesetz und der Schutz vor Diskriminierung durch algorithmische Entscheidungssysteme, 2023, S. 14.
[39] *Spiecker gen. Döhmann/Towfigh*, Das Allgemeine Gleichbehandlungsgesetz und der Schutz vor Diskriminierung durch algorithmische Entscheidungssysteme, 2023, S. 14; vgl. auch *Orwat*, Diskriminierungsrisiken durch Verwendung von Algorithmen, 2019, S. 3 ff.
[40] *Hoffmann-Riem,* AöR 2017, 1 (1); *Datenethikkommission der Bundesregierung (Hrsg.)*, Gutachten der Datenethikkommission, S. 34; zum Einfluss von Algorithmen auf Demokratie und Öffentlichkeit *Buchmann et al.*, Digitalisierung und Demokratie, passim.; *Spiecker gen. Döhmann* in *Kischel/Kube (Hrsg.)*, HStR[4], § 20, Rdnr. 12 ; *dies* in *Kahl/Ludwigs (Hrsg.),* HdbVerwR III, § 71 Rdnr. 6 f.
[41] Dazu statt vieler *Voßkuhle* in *Hoffmann-Riem/Schmidt-Aßmann/Voßkuhle:* Grundlagen des Verwaltungsrechts, § 1.

Thaler und *Cass R. Sunstein* haben hierzu den Ansatz des libertären Paternalismus entwickelt, der auf die Steuerung des Verhaltens der Bürger durch gezielte staatliche Maßnahmen gerichtet ist.[42] Dabei sollen in der Weiterentwicklung Nudges („Stupser" oder „Anstöße") dem Menschen dazu verhelfen, „bessere", da für den Einzelnen (oder auch die Gesellschaft) vorteilhaftere, Entscheidungen zu treffen, wobei ihre Entscheidungsfreiheit aber so wenig wie möglich eingeschränkt werden soll. Die Nudges sollen dabei eine sehr subtile Möglichkeit der Einflussnahme sein, ohne rechtliche Ge- und Verbote, Steuern oder Subventionen einzusetzen.[43] Sie setzen auf ein faktisches Element, mittels dessen eine Entscheidungsgestaltung dazu führt, dass die „richtigen" Entscheidungen getroffen werden, obwohl andere, „schlechtere" Entscheidungen weiterhin möglich sind.[44] Oftmals setzen sie gezielt auf psychologische Entscheidungsmuster. Damit soll ein Ausgleich geschaffen werden zwischen dem paternalistischen Element des Regierens und dem libertären Element, das größtmögliche Freiheit vor staatlichen Eingriffen verspricht.[45] Das Konzept ähnelt damit in Teilen dem in der deutschen (Verwaltungs-)Rechtswissenschaft weit verbreiteten Governance- oder auch steuerungswissenschaftlichen Ansatz, ist aber nicht mit ihm gleichzusetzen.[46] Während der libertäre Paternalismus auf Ermöglichung größtmöglicher Freiheit für den Einzelnen durch Steuerung abzielt, will der Governance-Ansatz Steuerung nutzen, um Ordnung herzustellen.[47] Die Steuerungswissenschaft bezieht zudem eine Vielzahl unterschiedlicher Instrumente in ihre Analyse ein, während der libertäre Paternalismus sich gezielt auf das Instrument der „Nudges" konzentriert. Gemeinsam ist beiden jedoch, dass sie versuchen, durch eine Bündelung staatlicher und nichtstaatlicher Steuerungsmechanismen gesellschaftliche Probleme effektiver zu lösen.[48] Die verschiedenen entscheidungsrelevanten Faktoren, die zum Nudging genutzt werden können, lassen sich – wie im Folgenden gezeigt wird – auch auf Big-Data-Analysen übertragen.[49] Nudges arbeitet – wie jede Form der Manipulation – damit, bestimmte Entscheidungsmuster abzubilden. Information, und Big Data insofern in besonderem Maße, kann dazu beitragen, diese Muster besser zu (er)kennen und ihre Wirkweise besser einzuschätzen. Abzugrenzen von Nudges im Sinne von *Thaler* und *Sunstein* sind Maßnahmen, die Unzulänglichkeiten des menschlichen Entscheidungsverhaltens bewusst ausnutzen, um Druck zu erzeugen. Dazu zählen im digitalen Bereich insbesondere „Dark Patterns", also manipulative Designmuster, die z.B. anhand von „Belohnungssystemen" wie Treuepunkten das Abmelden von Abonnements erschweren sollen oder die potenzielle Kunden durch Hinweise auf die faktisch nicht vorhandene Knappheit von Waren zum Kauf animieren sollen.[50]

2.1.1.1 Standardvorgaben („Default Rules"):

Eine erste mögliche Maßnahme zum Nudging ist die Gestaltung des Ausgangszustandes. Dass sich dadurch das Verhalten beeinflussen lässt, beruht auf der Annahme, dass viele Menschen die Tendenz haben, den Ist-Zustand überhöht zu bewerten und in der Konsequenz daran festzuhalten, auch wenn dies irrational ist (Status Quo Bias).[51] Ein Grund hierfür könnte sein, dass für den Menschen Verluste mehr Bedeutung haben als Gewinne

[42] Dazu grundlegend *Sunstein/Thaler*, The University of Chicago Law Review 2003, 1159-1202.
[43] *Thaler/Sunstein*: Nudge.
[44] Kritisch dazu vor allem *Gigerenzer/Todd/ABC Research* Group: Simple Heuristics That Make Us Smart, S. 27 ff.; in der Rechtswissenschaft außerdem *Wolff*, RW 2015, 194 (209 ff.); *Eidenmüller*, JZ 2011, 814 (820); *Kirchhof*, ZRP 2015, 136 (136); *Grunert,* in *Anderheiden/Bürkli/Heinig/Kirste/Seelmann (Hrsg.)*: Paternalismus und Recht, S. 9 ff.; *van Aaken* in *Kemmerer/Möllers/Steinbeis/Wagner (Hrsg.)*: Choice Architecture in Democracies, S. 172 ff.
[45] *Mende*, ZfPP 2016, 559 (563); *Eidenmüller*, JZ 2011, 814 (817 f.).
[46] Zum Governance-Ansatz grundlegend *Schuppert*, Governance und Rechtsetzung; *Hoffmann-Riem*, Die Governance-Perspektive in der rechtswissenschaftlichen Innovationsforschung; *Schuppert* in *Hoffmann-Riem/Schmidt-Aßmann/Voßkuhle (Hrsg.)*: Grundlagen des Verwaltungsrechts, § 16 Rdnr. 20 ff. – hier wird auch der Unterschied zwischen der Governance-Perspektive und dem steuerungstheoretischen Ansatz verdeutlicht.
[47] *Mende*, ZfPP 2016, 559 (568).
[48] *Mende*, ZfPP 2016, 559 (559); zu den verfassungsrechtlichen Fragen aus deutscher Perspektive *Gerg*, Nudging: Verfassungsrechtliche Maßstäbe für das hoheitliche Einwirken auf die innere Autonomie des Bürgers, 2019.
[49] In der Governance-Forschung wird insofern auch von „Internetgovernance" gesprochen, *Hoffmann-Riem*, Innovation und Recht – Recht und Innovation, 2016, S. 641 ff.
[50] *Gertz/Martini/Seelinger/Timko*, LTZ 2023, 3 (4).
[51] Dazu ausführlich *Kahnemann/Knetsch/Thaler,* The Journal of Economic Perspectives 1991, 193; *Thaler/Sunstein*: Nudge, S. 55 ff.

(sog. Verlustaversion).[52] Außerdem sind Menschen träge und entscheiden sich meist für den Weg des geringsten Widerstands und Aufwands.[53] Der Status Quo Bias kann daher durch eine gezielte Gestaltung der Standardvorgaben (Default Rules) ausgenutzt werden.[54] Dies lässt sich am Beispiel der Organspende verdeutlichen: Obwohl laut einer Befragung der Bundeszentrale für gesundheitliche Aufklärung[55] im Jahr 2022 84 % der Befragten dem Thema Organspende positiv gegenüberstanden, sind viele Menschen zu träge, die Entscheidung zur Ausstellung eines Organspendeausweises zu treffen und lassen sich daher stark von der Standardvorgabe lenken.[56] Unabhängig davon, ob man dies politisch und ethisch für richtig hält, ließe sich durch die Einführung einer Widerspruchslösung die Zahl der Organspenden in Deutschland daher vermutlich deutlich erhöhen.[57] Ein weiteres Beispiel für einen Bereich, in dem Standardvorgaben sinnvoll zur Entscheidungsarchitektur genutzt werden können, sind datenschutzfreundliche Voreinstellungen.[58]

Die Ergebnisse einer prädiktiven oder präskriptiven Big-Data-Analyse könnten bei der Umsetzung eines erwünschten Ergebniszustands Grundlage für die Gestaltung des Ausgangszustandes sein. Denkbar ist insofern etwa, das Ergebnis der Datenverarbeitung nicht nur zur Gestaltung gesellschaftlicher, sondern ebenso für individualisierte Standardvorgaben zu nutzen. So könnte beispielsweise mithilfe eines Algorithmus die Wahrscheinlichkeit, mit der eine konkrete Person in Organspenden einwilligt oder nicht, errechnet und die Standardvorgabe – womöglich sogar eine voreingestellte Auswahl bestimmter Organe, bei denen aufgrund der Datenauswertung eine Zustimmung wahrscheinlich ist – dementsprechend angepasst werden.[59] Aus Platzgründen verbleibt es in diesem allgemeinen Rahmen bei diesem Beispiel.

2.1.1.2 Soziale Normen

Darüber hinaus lassen sich Entscheidungsprozesse durch sog. soziale Nudges lenken. Diese Möglichkeit der Einflussnahme geht zurück auf die Erkenntnis, dass sich Menschen aufgrund der Herausbildung sozialer Normen und sozialen Drucks stark durch ihr Umfeld beeinflussen lassen.[60] Eine Möglichkeit des sozialen Nudgings besteht darin, auf das Mehrheitsverhalten aufmerksam zu machen. So lässt sich beispielsweise der Energieverbrauch in Privathaushalten deutlich senken, wenn – unter Nennung realer Verbrauchszahlen – auf den deutlich geringeren Verbrauch der Nachbarn aufmerksam gemacht wird.[61] Hier bietet die Auswertung von Big Data großes Potenzial: Durch die große Zahl an Daten, die ausgewertet werden, lässt sich das Mehrheitsverhalten sowohl in der Vergangenheit als auch für die Zukunft mit hoher Wahrscheinlichkeit voraussagen und so treffsicher der „richtige" Anreiz geschaffen werden.

---

[52] *Wolff,* RW 2015, 194 (199); *Heeren*: Neuronale Grundlagen der Verlustaversion, S. 4; *Maturana*, DÖV 2022, 941 (944); *Klöhn*: Kapitalmarkt, Spekulation und Behavioral Finance, S. 95. *Wolff,* RW 2015, 194 (199).
[53] *Thaler/Sunstein*: Nudge, S. 123.
[54] *Thaler/Sunstein*: Nudge, S. 123 ff.
[55] *Bundesregierung*, Umfrage: Mehrheit steht Organspende positiv gegenüber, https://www.bundesregierung.de/breg-de/suche/tag-der-organspende-2047324.
[56] *Sunstein/Thaler*, The University of Chicago Law Review, 2003, 1159 (1192); vgl. *Johnson/Goldstein*, SCIENCE 2003, Vol. 302, 1338 (1338).
[57] Nach § 3 Transplantationsgesetz (TGP) ist die Entnahme von Organen derzeit nur zulässig, wenn der Spender in die Entnahme eingewilligt hat (Zustimmungslösung); nach § 4 TPG ist es unter Umständen aber möglich, dass ein naher Angehöriger in die Entnahme einwilligt.
[58] Hintergrund für das Bedürfnis nach solchen „Privacy Nudges" ist das sog. „Privacy Paradox", das beschreibt, dass vielen Menschen ihre Privatsphäre zwar wichtig ist, sie aber trotzdem nicht datenschutzfreundlich handeln und stattdessen eine Vielzahl persönlicher Informationen über sich teilen, zum Privacy Paradox *Barth/de* Jong, Telemedicus and Informatics 2017, 1038 ff. Privacy Nudges sollen diesem Problem Abhilfe verschaffen, dazu: *Barev et al.*, Systematisches Design digitaler Privacy Nudges; *Schomberg et al.*, DuD 2019, 774 ff.; *Schöbel et al.* in *Friedewald/Kreutzer/Hansen (Hrsg.)*, Selbstbestimmung, Privatheit und Datenschutz, S. 369 ff.
[59] Das riefe freilich erhebliche Bedenken im Hinblick auf den Gleichheitsgrundsatz aus Art. 3 GG hervor, näheres dazu unten S.21.
[60] *Thaler/Sunstein*: Nudge, S. 79 ff.; *Wolff,* RW 2015, 194 (201).
[61] Verstärkt wurde dieser Effekt noch, wenn das Schreiben zur Mitteilung des Energieverbrauchs mit einer visuellen Bewertung des individuellen Energieverbrauchs in Form eines traurigen oder lachenden Smileys versehen war, *Thaler/Sunstein*: Nudge, S. 101 f.

2.1.1.3 Offenlegung von Information

Eine weitere Form des Nudgings besteht in der gezielten Bereitstellung von Information und Salienz.[62] Da der Mensch Informationen nur begrenzt speichern kann und nicht alle Informationen verstehen und in seine Entscheidungen mit einbeziehen kann, können gezielt eingesetzte Informationen zu einer besseren Erreichung des gewünschten Handelns führen.[63] Außerdem schätzen Menschen Wahrscheinlichkeiten über den Eintritt eines Ereignisses häufig falsch ein, da sie sich zur Urteilsfindung an Beispielen orientieren, die in der Erinnerung wegen jüngst zurückliegender Aktualisierungen besonders präsent sind – etwa persönliche Erfahrungen oder Medienberichte (sog. „Verfügbarkeitsheuristik").[64] Dieser Urteilsfehler kann durch gezieltes Einsetzen von Information zur Verhaltensbeeinflussung genutzt werden. Auch die Art der Informationsvermittlung ist von Bedeutung: so kann die vereinfachte Darstellung komplexer Sachverhalte eine Entscheidung erheblich „erleichtern".[65] Dabei ist aber freilich nicht jede Information ein Nudge. Der Staat stellt in vielfältiger Weise Informationen bereit und setzt diese auch gezielt zur Steuerung ein.[66] Im Gegensatz zu sonstigem staatlichen Informationshandeln stellt die Bereitstellung von Informationen einen Nudge dar, wenn sie – wenn auch nur subtil – dazu dienen soll, die Entscheidungsarchitektur zu beeinflussen.[67] Beispiele für informationelle Nudges sind Warnhinweise auf Tabakprodukten oder der in Deutschland seit 2020 auf Lebensmitteln abgedruckte „Nutri-Score".[68]

Wie bereits gezeigt, lassen sich mithilfe der algorithmisierten Auswertung großer, realer Datensätze neue Muster und Querverbindungen herstellen – und damit neue Informationen generieren, die einem menschlichen Betrachter nicht aufgefallen wären. Diese Informationen können dann zur Steuerung von Entscheidungsprozessen gezielt eingesetzt werden. Auch können Textverarbeitungs-KIs wie ChatGPT eingesetzt werden, um komplizierte Informationen vereinfacht darzustellen.[69]

2.1.1.4 Personalisierung

Ein weiteres Mittel zur Steuerung von Verhalten sind Formen der personalisierten Adressierung.[70] Eine individuelle und personalisierte Ansprache – von der bloßen Verwendung des Namens bis zur Einbeziehung weiterer Daten wie Wohnort, Alter, Verhalten oder Interessen – führt dazu, dass die Aufmerksamkeit grundsätzlich steigt und der Einzelne eher zum Handeln motiviert wird. Außerdem fühlt sich der persönlich Angesprochene in der Regel weniger anonym, was eine disziplinierende Wirkung hat.[71] Nach dem Ansatz des libertären Paternalismus wäre Personalisierung als sozialer Nudge einzuordnen. Werden Nudges individuell an das Persönlichkeitsprofil einer Person angepasst, das aufgrund einer Auswertung großer Datensätze prognostiziert wurde, ist auch von „Hypernudging" die

---

[62] *Thaler/Sunstein*: Nudge, S. 260 ff.; *Baer:* Staatliche Steuerung durch Nudging, S. 109 ff; *Sunstein*, Yale Journal on Regulation 2015, 413 (429).
[63] *Meier*: Verhaltenswissenschaftlich inspiriertes Verwaltungshandeln, S. 42.
[64] *Kahnemann*: Schnelles Denken, langsames Denken, S. 164 ff.; kritisch *Gigerenzer/Todd/The ABC Research Group*: Simple Heuristics That Make Us Smart, passim.
[65] *Sunstein*: Simpler, S. 90 ff.; *Kronenberger*: Nudging als Steuerungsinstrument des Rechts, S. 68; zur Darstellung von Informationen zu Impfrisiken, um eine informierte Impfentscheidung zu ermöglichen, *Spiecker* in *Engel/Englerth/Lüdemann/dies. (Hrsg.)*: Recht und Verhalten, 133 (144 ff.).
[66] *Baer*, Staatliche Steuerung durch Nudging, S. 34.
[67] *Baer*, Staatliche Steuerung durch Nudging, S. 38.
[68] Vgl. *Baer*: Staatliche Steuerung durch Nudging, S. 109; gem. § 6 I TabakerzeugnisG, §§ 12-17 TabakerzeugnisVO sind die Verpackungen von Tabakerzeugnisse mit gesundheitsbezogenen Warnhinweisen zu versehen; der Nutri-Score soll vereinfacht Aufschluss über die Nähr- und Gesundheitswerte von Lebensmitteln geben und kann von den Unternehmen freiwillig auf Produkten abgedruckt werden, s. dazu *Bundesministerium für Ernährung und Landwirtschaft*, Nutri-Score, https://www.bmel.de/DE/themen/ernaehrung/lebensmittel-kennzeichnung/freiwillige-angaben-und-label/nutri-score/nutri-score_node.html.
[69] Vgl. *Breher/Lehmann,* Vordenker zu ChatGPT, https://www.tagesspiegel.de/chancen-und-gefahren-der-kunstlichen-intelligenz-das-ist-schon-ziemlich-revolutionar-9558341.html.
[70] S. dazu z.B. *Spiecker gen. Döhmann*, VVDStRL 2017, 10 (38) mwN.
[71] *Wolff,* RW 2015, 194 (202); *Kronenberger*: Nudging als Steuerungsinstrument des Rechts, S. 67.

Rede.[72] Doch kann Beeinflussung durch Personalisierung auch noch einen Schritt weiter gehen und damit eine eigenständige Kategorie der Verhaltensbeeinflussung bilden: Zeigt uns ein Algorithmus von vorneherein nur noch bestimmte, personalisierte Inhalte an, kann es passieren, dass der Einzelne faktisch gar keine Auswahlmöglichkeit zwischen verschiedenen Optionen hat und ihm nur noch bestimmte Optionen zugänglich sind bzw. die Wahl anderer Optionen unverhältnismäßig aufwendig wird.[73]

Ein Bereich, in dem Personalisierung bereits gezielt zur Beeinflussung menschlichen Verhaltens eingesetzt wird, ist das Microtargeting.[74] Dabei werden große Datensätze von einer KI ausgewertet, um Personen – etwa Nutzer sozialer Netzwerke als potenzielle Wähler oder Kunden – in unterschiedliche Zielgruppen zu unterteilen und der jeweiligen Personengruppe gezielt und personalisiert bestimmte Inhalte und Informationen zuspielen zu können. Nachdem Microtargeting sich im kommerziellen Bereich bereits als Werbestrategie etabliert hat, ist es spätestens seit den US-Wahlen und dem Brexit-Referendum 2016 auch in die Politik vorgedrungen. In einer noch gesteigerten Form der Personalisierung können Big-Data-Auswertungen außerdem dazu genutzt werden, umfassende Persönlichkeitsprofile zu erstellen und den Einzelnen psychometrisch zu vermessen. Diese können dann z.B. die Grundlage für das in Europa besonders kritisch betrachtete „Social Scoring"[75] bilden.[76]

2.1.2 Chilling Effects

Gerade wenn sie im Bereich der Polizei- und Ermittlungsarbeit eingesetzt wird, kann die Big-Data-Auswertung zudem ein Überwachungsgefühl bei den Betroffenen verursachen, das dazu führt, dass sie ihr Verhalten vorausschauend im Hinblick auf vermeintliche Konsequenzen anpassen.[77] Solche Verhaltensanpassungen aufgrund des Gefühls, überwacht zu werden, werden als „chilling effects" oder „Abschreckungseffekte" bezeichnet.[78] Allerdings bedeutet das nicht umgekehrt, dass das Vertrauen, nicht überwacht zu werden, unmittelbar zurückkehrt, wenn solche Datenauswertungen zurückgefahren werden; mit anderen Worten: Überwachungsgefühl und tatsächlicher Überwachungsgrad korrelieren häufig nicht unmittelbar miteinander.

2.1.3 Automation Bias

Schließlich besteht besonders bei automatisierter Entscheidungsfindung eine erhöhte Gefahr, dass die Personen, die große, reale Datenverarbeitungssysteme anwenden, die Resultate der Datenverarbeitung nicht einer eigenen Plausibilitätskontrolle unterziehen und sich mehr oder weniger „blind" ihrem Vertrauen in die Technik hingeben. Diese Tendenz, sich maschinellen Empfehlungen vorbehaltslos hinzugeben und ihnen mehr zu vertrauen als menschlichen Entscheidungen, wird als Automation Bias bezeichnet.[79] Dies ist vermutlich u.a. darauf zurückzuführen, dass bei maschinellen Urteilen häufig eine Objektivität

[72] Dazu *Yeung*, Information, Communication & Society, 118-136.
[73] *Spiecker gen. Döhmann*, VVDStRL 2017, 10 (38).
[74] *Kelber/Leopold* in *Spiecker gen. Döhmann/Westland/Campos (Hrsg.)*: Demokratie und Öffentlichkeit im 21. Jahrhundert, S. 160 ff.; *Kind/Weide*: Microtargeting, 2017.
[75] Wobei hierfür im KI-VO-E ein weitreichendes Verbot geplant ist, s. die Erläuterungen zu Teil II der Verordnung und Art. 5 KI-VO-E.
[76] *Spiecker gen. Döhmann* in *Kahl/Ludwigs (Hrsg.)*, HdbVerwR III, § 71 Rdnr. 8; *dies.* in *Kube/Kischel (Hrsg.)*, HStR[4], § 20 Rdnr. 12 .
[77] So auch schon das BVerfG im Volkszählungsurteil, BVerfGE 65, 1 (42 f.): „Damit haben sich in einer bisher unbekannten Weise die Möglichkeiten einer Einsicht- und Einflußnahme erweitert, welche auf das Verhalten des Einzelnen schon durch den psychischen Druck öffentlicher Anteilnahme einzuwirken vermögen. […] Wer unsicher ist, ob abweichende Verhaltensweisen jederzeit notiert und als Information dauerhaft gespeichert, verwendet oder weitergegeben werden, wird versuchen, nicht durch solche Verhaltensweisen aufzufallen."
[78] Dazu schon ausführlich *Büscher et al.:* Verhaltensbeeinflussung durch Beobachtung?, S. 18 ff.; zu Abschreckungseffekten *Staben*: Der Abschreckungseffekt auf die Grundrechtsausübung; *Schwabenbauer*: Heimliche Grundrechtseingriffe, S. 140 ff.; *Assion* in *Telemedicus e.V. (Hrsg.)*: Überwachung und Recht, S. 31 ff.; *Oermann/Staben*: Der Staat 2013, 630; S. auch *Spiecker* in *Kischel/Kube (Hrsg.)*, HStR[4], § 20 Rdnr. 32.
[79] *Deutscher Ethikrat*: Mensch und Maschine, S. 142; *Guijarro Santos,* ZfDR 2023, 23 (28); *Raue*, ZUM 2023, 160 (167); *Skitka/Mosier/Burdick,* Int. J. Human-Computer Studies 1999, 991 (992).

unterstellt wird.[80] Zumindest unbewusst wird in der Folge Verantwortung leichter an maschinelle Entscheidungssysteme delegiert.[81] Während die vorangehend dargestellten Steuerungsformen eher aus der Perspektive des Bürgers als Adressat staatlicher Maßnahmen gedacht sind, bezieht sich diese eher auf Szenarien, in denen sich der Staat zur Entscheidungsfindung großer Datenverarbeitungssysteme bedient und das Verhalten des menschlichen Entscheiders, der die Ergebnisse einer Big-Data-Analyse auswertet, beeinflusst wird. Verzerrungen können dabei nicht nur bei Entscheidungen zum Tragen kommen, die gänzlich ohne menschliches Hinzutun getroffen werden, sondern bereits bei einer Entscheidungsunterstützung.[82] Man kann in § 22 DSGVO auch eine Reaktion auf diesen Automation Bias sehen.

2.2. Grundrechtsrelevanz der Verhaltensbeeinflussung durch Big-Data-Analysen

Staatliches Handeln findet seine äußersten Grenzen in den Bestimmungen des Grundgesetzes, insbesondere in den Grundrechten, die die Grenzen für Eingriffe in den Freiheitsbereich der Bürger abstecken und an die die gesamte Staatsgewalt nach Art. 1 III GG unmittelbar gebunden ist. Verarbeitet der Staat Daten der Bürger und macht diese Verarbeitungen zur Grundlage für verhaltenssteuernde Maßnahmen, so ist dieses Vorgehen an den Grundrechten zu messen. In Betracht kommen dabei insbesondere die informationelle Selbstbestimmung, die allgemeine Handlungsfreiheit als Recht zu freiem Entscheiden sowie der allgemeine Gleichbehandlungsgrundsatz. Grundsätzlich können aber – je nach Einzelfall – alle Grundrechte betroffen sein. Neben den im Folgenden behandelten Grundrechten kommen insbesondere noch die Versammlungsfreiheit und das Telekommunikationsgeheimnis in Betracht. Eine umfassende Behandlung würde den zur Verfügung stehenden Rahmen indes sprengen. Die hiesige Auswahl konzentriert sich daher auf grundlegende Grundrechte und kann auch diese nur pointiert skizzieren.

2.2.1 Informationelle Selbstbestimmung

Das Grundrecht auf informationelle Selbstbestimmung als Ausprägung des allgemeinen Persönlichkeitsrechts (Art. 2 I i.V.m. Art. 1 I GG) ist auf den Schutz personenbezogener Daten gerichtet und garantiert das Recht des Einzelnen, „grundsätzlich selbst zu entscheiden, wann und innerhalb welcher Grenzen persönliche Lebenssachverhalte offenbart werden“[83]. Europarechtlich ist das Grundrecht auf Schutz personenbezogener Daten in Art. 8 (und tlw. in Art. 7) GRCh sowie Art. 8 EMRK verankert.[84] Der Schutzbereich des Grundrechts erstreckt sich zwar grundsätzlich auf jede Form der Datenverarbeitung;[85] durch die gesteigerten technischen Möglichkeiten erfährt der Schutz vor computerbasierter Datenverarbeitung aber einen besonderen Bedeutungszuwachs.[86] Aufgrund der umfassenden digitalen Vernetzung und Vermessung unserer gesamten Lebenswelt kann der Einzelne häufig gar nicht mehr überblicken, wann und wo personenbezogene Informationen gesammelt werden. Gerade angesichts der Möglichkeit, Persönlichkeitsprofile zu er-

[80] *Deutscher Ethikrat*: Mensch und Maschine, S. 135 f.; *Horstmann*, ZD-Aktuell 2020, 07047.
[81] *Deutscher Ethikrat*: Mensch und Maschine, S. 136; *Skitka/Mosier/Burdick,* Int. J. Human-Computer Studies 1999, 991 (1001).
[82] *Deutscher Ethikrat*: Mensch und Maschine, S. 226; vgl. *Sesing/Tschech*, MMR 2022, 24 (28).
[83] BVerfGE 65, 1 (42) – „Volkszählung“. BVerfGE 65, 1 (42); konkretisiert durch: BVerfGE 113, 348 – „Telekommunikationsüberwachung“; 115, 320 – „Rasterfahndung“; 120, 274 – „Online-Durchsuchung“; 125, 260 – „Vorratsdatenspeicherung“; in der Literatur zum Recht auf informationelle Selbstbestimmung: *Kunig/Kämmerer* in: v. Münch/Kunig, GG, Art. 2, Rn. 75 ff.; *Starck* in: v. Mangoldt/Klein/Starck, GG, Art. 2, Rn. 114 ff.; *Rixen* in: Sachs, GG, Art. 2, Rn. 72 ff.; *Dreier* in ders., GG, Art. 2 Rn. 79 ff.; *Di Fabio* in Dürig/Herzog/Scholz, GG, Art. 2, Rn. 173 ff.
[84] *Schiedermair* in *Simitis/Hornung/Spiecker gen. Döhmann (Hrsg.)*, DSGVO/BDSG, Einleitung, Rdnr. 161; *Kingreen* in *Calliess/Ruffert*, EUV/AEUV, Art. 8 GRCh Rdnr. 5, 10.
[85] Also insbesondere auch die manuelle Datenverarbeitung, BVerfGE 78, 77 (84); *Di Fabio* in *Dürig/Herzog/Scholz*, GG, Art. 2 Rdnr. 176.
[86] *Dreier* in *ders.*, GG, Art. 2 Rdnr. 79; *Kunig/Kämmerer* in *v. Münch/Kunig*, GG, Art. 2 Rdnr. 75.

stellen, muss der Einzelne davor geschützt werden, zum bloßen Objekt des Staates degradiert zu werden[87]. Da Datenverarbeitungen zunehmend nicht mehr nur durch den Staat, sondern vor allem auch durch Private erfolgt, kommt dem Staat insofern eine – noch in den Details näher zu bestimmende – Schutzpflicht zu.[88]

Im Rahmen der DSGVO gilt der Grundsatz der Datenminimierung (Art. 5 Abs. 1 lit. c DSGVO)[89], wonach so wenig Daten mit Personenbezug wie möglich verarbeitet werden dürfen. Wenn dementsprechend Big-Data-Anwendungen ausschließlich anonymisierte Daten verwenden, ist in diesen Fällen der Anwendungsbereich der DSGVO gem. Art. 2 Abs. 1 gar nicht erst eröffnet. Das bedeutet aber nicht, dass solche Verarbeitungen im Hinblick auf das Grundrecht auf informationelle Selbstbestimmung gänzlich irrelevant sind – vor allem, da meist nicht auszuschließen ist, dass der Personenbezug der Daten wiederhergestellt werden kann.[90] Werden Big-Data-Analysen jedoch dazu eingesetzt, in irgendeiner Form zu personalisieren, etwa beim Profiling, Scoring oder Microtargeting, werden dazu (zumindest auch) Daten mit – neu hergestelltem – Personenbezug verwendet. Daher bedarf es auch eines mindestens dreischrittigen Kontrollansatzes.[91] Ein Regulierungsansatz muss danach im ersten Schritt zunächst grundlegend den Erwerb, die Nutzung, Speicherung und Weiterverarbeitung von Daten kontrollieren. Des Weiteren muss die Auswertung der Daten begleitet und strukturiert werden. Im dritten und letzten Schritt müssen die Ergebnisse und Folgen der Datenauswertung regulativ erfasst werden.[92]

Entscheidet das Datenverarbeitungssystem automatisiert, stellen sich weitere Probleme: Auf etwaige Grundrechtseingriffe wirkt bei automatisierten Entscheidungen verstärkend, dass die Selbstentwicklung der eingesetzten algorithmischen Systeme eine Erklärbarkeit des Ergebnisses häufig unmöglich macht.[93] Insgesamt können die lange Speicherdauer, die dezentrale Abrufbarkeit und leichte Verknüpfbarkeit von Daten durch die Digitalisierung sowie die Vielzahl von Einzeleingriffen, die z.B. auch durch Zweckänderungen bei großen, realen Datenverarbeitungssystemen entstehen, die Gesamtintensität der Grundrechtseingriffe erhöhen.[94] Außerdem können mit den Chilling Effects Eingriffe in die informationelle Selbstbestimmung auch Auswirkungen auf die Ausübung sämtlicher anderer Freiheitsrechte haben. Selbst wenn solche Verhaltensbeeinflussungen für sich genommen keinen Eingriffscharakter haben, müssen sie dennoch in die Gesamtbetrachtung mit einbezogen werden.[95]

---

[87] *Di Fabio* in *Dürig/Herzog/Scholz*, GG, Art. 2 Rdnr. 173; BVerfGE 65, 1 (42) – „Volkszählung".
[88] *Dreier* in *ders.*, GG, Art. 2 Rdnr. 81; zu Defiziten des Grundrechtsschutzes im Internet *Schliesky/Hoffmann/Luch/Schulz/Borchers*: Schutzpflichten und Drittwirkung im Internet, S. 119 ff.; *Muckel,* VVDStRL 2019, 79 (245); *Peuker*: Verfassungswandel durch Digitalisierung, S. 295 ff.; *Hoffmann/Luch/Schulz/Borchers*: Die digitale Dimension der Grundrechte, S. 64 ff.
[89] Zum Grundsatz der Datenminimierung und der daraus folgenden Erforderlichkeit der Reduzierung des Personenbezugs von Daten *Roßnagel* in *Simitis/Hornung/Spiecker* gen. Döhmann *(Hrsg.)*, DSGVO/BDSG, Art. 5 Rdnr. 116 ff; *Pötters* in *Gola/Heckmann (Hrsg.)*, DSGVO/BDSG, Art. 5 Rdnr. 22 ff.
[90] Vgl. *Weichert,* ZD 2013, 152 (254); *Friele/Bröckerhoff/Fröhlich/Spiecker gen. Döhmann/Woopen*, Bundesgesundheitsblatt 63 (2020), 741 (746); näheres zur Anonymisierung von Daten und dem Risiko der Wiederherstellung des Personenbezugs unten, S. 28.
[91] S. dazu *Spiecker gen. Döhmann* in *Kischel/Kube (Hrsg.)*, HStR[4], § 20 Rdnr. 56 ; vgl. *Paal/Hennemann*, NJW 2017, 1697 (1700); *Hackenberg* in *Hoeren/Sieber/Holznagel (Hrsg.),* HdbMultimediaR, Teil 15.2 Rdnr. 13 ff.
[92] *Spiecker gen. Döhmann* in *Kischel/Kube (Hrsg.)*, HStR[4] 2023, § 20, Rn. 56.
[93] *Spiecker gen. Döhmann* in *Kahl/Ludwigs (Hrsg.)*, HdbVerwR III, § 71 Rdnr. 16 mwN; *Zech*, ZfPW 2019, 198 (205); *Gausling*, ZD 2019, 335 (335).
[94] *Spiecker gen. Döhmann* in *Kahl/Ludwigs (Hrsg.)*, HdbVerwR III, § 71 Rdnr. 59 f.; *dies.* in *Kischel/Kube (Hrsg.)*, HStR[4], § 20 Rdnr. 10 f. ; *Ruschemeier*: Der additive Grundrechtseingriff, passim.
[95] Vgl. (allerdings ausschließlich in Bezug auf Überwachungsmaßnahmen) BVerfGE 125, 260 (324) – „Vorratsdatenspeicherung"; *Roßnagel*, NJW 2010, 1238 (1240). Der Ansatz einer Überwachungsgesamtrechnung findet sich außerdem auch im Koalitionsvertrag der Ampelregierung wieder, SPD/Die Grünen/FDP, Mehr Fortschritt wagen. Koalitionsvertrag zwischen SPD, Bündnis 90/Die Grünen und FDP, S. 108. Zudem wird in aktuellen Forschungsansätzen wie dem periodischen Überwachungsbarometer des MPI zur Erforschung von Kriminalität, Sicherheit und Recht die praktische Umsetzbarkeit untersucht.

2.2.2 Allgemeine Handlungsfreiheit

Mit Blick auf die staatliche Steuerung kommt außerdem eine Beeinträchtigung der allgemeinen Handlungsfreiheit (Art. 2 I GG) in Betracht, sofern der Einzelne keine freie Entscheidung mehr treffen kann. Ob sich dem Grundgesetz, insbesondere Art. 2 I GG, ein Recht auf freies Entscheiden entnehmen lässt, hängt aber maßgeblich von dem Grundverständnis von Freiheit, Autonomie und Selbstbestimmung ab, das dem Grundgesetz innewohnt bzw. bei seiner Auslegung als Vorverständnis zugrunde gelegt wird.[96] Noch immer ist dabei das Autonomieverständnis *Kants* von zentraler Bedeutung.[97] Die Autonomie des Willens besteht nach *Kant* darin, „nicht anders zu wählen als so, daß die Maximen seiner Wahl in demselben Wollen zugleich als allgemeines Gesetz mit begriffen seien"[98] – ist also im Sinne von Selbstgesetzgebung zu verstehen.[99] Man könnte Autonomie danach folglich als negative Freiheit interpretieren, das heißt: als die Absenz externer Beschränkungen.[100] Für das BVerfG gehört sogar zur Menschenwürde als *der* Fundamentalnorm des Grundgesetzes, „daß der Mensch über sich selbst verfügen und sein Schicksal eigenverantwortlich gestalten kann"[101]. Verantwortung kann der Mensch aber nur übernehmen, wenn er auch *selbst* entscheidet. Voraussetzung einer solchen Zurechnung ist Freiheit. Dem Gedanken von Autonomie und Selbstbestimmung wird daher auch mit den Freiheitsrechten, vor allem mit der allgemeinen Handlungs- als Basisfreiheit aus Art. 2 I GG Ausdruck verliehen.[102]

Die Grundrechte des Grundgesetzes werden aber schon seit Langem nicht allein in ihrer Abwehrdimension verstanden[103], sondern, wie es in einer prominenten Formulierung heißt, zumindest auch als „Optimierungsgebote"[104]. Konzepte wie das Nudging entsprechen dabei gerade nicht dem Idealbild eines autonomen Entscheidungsprozesses. Wenngleich Nudges von ihrer Grundidee nicht in Freiheitsrechte eingreifen, sondern freies Entscheiden erst ermöglichen sollen, ist dennoch zumindest der Schutzbereich der allgemeinen Handlungsfreiheit betroffen;[105] ob und wie intensiv ein Eingriff vorliegt, ist dann eine Frage des Einzelfalls. Gerade wenn aber durch die Datenverarbeitung Eingriffe in die informationelle Selbstbestimmung hinzukommen, wird steuerndes Handeln von Seiten des Staates regelmäßig rechtfertigungsbedürftig sein. Liegen gute Gründe für den Einsatz eines Nudges vor, wird diese Rechtfertigung auch gelingen können. Je extensiver der Staat sich dieses verhaltenssteuernden Mittels bedient, desto höher müssen indes die qualitativen Anforderungen an die rechtfertigenden Gründe im Einzelfall ausfallen, um in der Gesamtschau nennenswerte Bereiche echter Autonomie zu gewährleisten. Nudging ist insofern zwar unter Umständen eine freiheitsschonende Beeinflussungsform, kann aber im Einzelfall durchaus auch verfassungswidrig sein. Nicht unterschätzt werden darf eben die indirekte und damit für den Betroffenen kaum erkennbare Wirkung.

[96] So auch *Baer*: Staatliche Steuerung durch Nudging, S. 151; *Honer*, DÖV 2019, 940 (944); vgl. *Hufen*, JuS 2020, 193 (197).
[97] Insbesondere bei der Auslegung der Menschenwürdegarantie aus Art. 1 I GG, deren zentraler Kern die Gewährleistung von Autonomie bildet, s. nur *Herdegen* in *Dürig/Herzog/Scholz*, GG, Art. 1 Rdnr. 12.
[98] *Kant*: Grundlegung zur Metaphysik der Sitten, zitiert nach *Weischedel (Hrsg.)*, S. 74.
[99] *Sacksofsky* KJ 2021, 47 (47); *Pfordten* in *Görres-Gesellschaft (Hrsg.)*, Staatslexikon, Menschenwürde.
[100] So auch *Sacksofsky* KJ 2021, 47 (48).
[101] BVerfGE 49, 286 (298) – „Transsexuelle".
[102] Vgl. nur *Di Fabio* in *Dürig/Herzog/Scholz*, GG, Art. 2 Rdnr. 13: Art. 2 I GG schützt den „Selbstentwurf des Menschen nach seinem Willen", zum Teil wird auch vertreten, allen Grundrechten des GG wohne ein absolut geschützter Kern der Menschenwürde inne *Denninger* JZ 1998, 1129 (1134); *Hilgendorf*: Menschenwürde und Demütigung, S. 149; *Huber* in *v. Mangoldt/Klein/Starck*, GG, Art. 19 Rdnr. 126; kritisch gegenüber einem Menschenwürdekern aber *Dreier* in *ders.*, GG Art. 1 Rdnr. 163; *Isensee* in *Merten/Papier*, HGR IV, § 87 Rdnr. 121 ff.
[103] BVerfGE 7, 198 – „Lüth"; *Grabenwarter* in *Dürig/Herzog/Scholz*, GG, Art. 5 Rdnr. 106 f.
[104] *Alexy*, Theorie der Grundrechte, S. 71 ff.
[105] Auch *Baer*: Staatliche Steuerung durch Nudging, S. 182 sieht die Entscheidungsfreiheit als vom Schutzbereich der Grundrechte geschützt und siedelt einen gewissen Integritätsschutz sowie den Schutz bestimmter Einstellungen im Einzelfall beim APR an, während die Willensentschließungsfreiheit von Art. 2 I GG geschützt sein soll; vgl. *Honer*, DÖV 2019, 940 (946); *Purnhagen/Reisch*, ZEuP 2016, 629 (646).

2.2.3 Ungleichbehandlung

Schließlich könnten algorithmengesteuerte Verhaltensbeeinflussungen durch Auswertung großer Datensätze im öffentlichen Sektor Ungleichbehandlungen i.S.v. Art. 3 I bzw. III GG hervorrufen. Häufig wird zwar damit geworben, automatisiertes Entscheiden ermögliche objektivere Maßstäbe, da das menschlichem Entscheiden innewohnende Element der Willkür ausgeschaltet werden könne;[106] empirische Untersuchungen zeigen jedoch, dass algorithmisches Entscheiden und künstliche Intelligenz faktisch in vielerlei Hinsicht diskriminierungsanfällig sind[107].

Ungleichbehandlungen können dabei auf unterschiedliche Weise entstehen. Zum einen können die persönlichen Überzeugungen und Ansichten ihrer Schöpfer in die Programmierung von Algorithmen einfließen.[108] Zwar könnte man meinen, dass es einfach zu erkennen wäre, ob dabei etwa verbotene Differenzierungsmerkmale nach Art. 3 III GG einbezogen wurden; da bei vielen algorithmischen Systemen der Programmcode – bzw. die dem Programmcode zugrundeliegenden Algorithmen – und die Trainingsdaten jedoch nicht offenliegen und zudem allenfalls Experten, zum Teil aber nicht einmal diese, in der Lage sind, diesen Code zu verstehen oder die Verzerrungsanfälligkeit der Trainingsdaten zu analysieren, gestaltet sich der Nachweis einer Diskriminierung in der Praxis komplizierter als es zunächst scheinen mag.[109] Bei selbstlernenden Systemen besteht zudem die Gefahr, dass die künstliche Intelligenz sich mit der Zeit selbst diskriminierendes Verhalten beibringt und sogar die Entwickler des Systems nicht mehr nachvollziehen können, welche Kriterien Maßstab der Entscheidung geworden sind.[110] Bei automatisiertem Entscheiden drohen Diskriminierungen durch vorherige Standardisierung und Gruppenbildung zudem verfestigt zu werden.[111]

Diskriminierungen durch Algorithmen können außerdem Beeinträchtigungen der allgemeinen Handlungsfreiheit zur Folge haben: Durch die Bewertung bestimmter Merkmale einer Person entsteht ein bestimmtes Bild. Die Betroffenen werden daher häufig mit einer fremdbestimmten Konstruktion ihrer Identität konfrontiert, die indirekt zu Verhaltensbeeinflussungen führen kann, indem die betroffene Person – ähnlich den Chilling Effects – ihre Selbstdarstellung anpasst.[112]

2.3 Anwendungsbereiche für algorithmengesteuerte Verhaltensbeeinflussung im öffentlichen Sektor

Im Folgenden werden nun beispielhaft einige besonders relevante Bereiche des öffentlichen Sektors aufgezeigt, in denen Verhaltensbeeinflussungen durch Big-Data-Analysen eine Rolle spielen oder in Zukunft Bedeutung erlangen könnten.

---

[106] *Wischmeyer,* AöR 2018, 1 (26); vgl. *Ernst*, JZ 2017, 1026 (1029).
[107] *Spiecker gen. Döhmann* in *Kahl/Ludwigs (Hrsg.)*, HdbVerwR III, § 71 Rdnr. 14; *Fröhlich/Spiecker gen. Döhmann*, Können Algorithmen diskriminieren?, https://verfassungsblog.de/koennen-algorithmen-diskriminieren/; zu den empirischen Untersuchungen *Bozdag*, Ethics and Information Technology, Vol. 15, Issue 3, 2013, (209); *Crawford/Calo*, Nature News 2016 vol. 438, 311 ff.
[108] *Martini,* JZ 2017, 1017 (1018); *Kischel* in *Epping/Hillgruber*, BeckOK GG, Art. 3 Rdnr. 218d.2.
[109] *Spiecker gen. Döhmann/Towfigh*, Das Allgemeine Gleichbehandlungsgesetz und der Schutz vor Diskriminierung durch algorithmische Entscheidungssysteme, 2023, S. 28 ff.; *Grünberger*, ZRP 2021, 232 (233).
[110] *Kischel* in *Epping/Hillgruber*, BeckOK GG, Art. 3 Rdnr. 218c; *Lauscher/Legner*, ZfDR 2022, 367 (375 f.); *Wischmeyer*, AöR 2018, 1 (3).
[111] *Spiecker gen. Döhmann* in *Kahl/Ludwigs (Hrsg.)*, HdbVerwR III, § 71 Rdnr. 14; *Fröhlich/Spiecker gen. Döhmann*, Können Algorithmen diskriminieren?, https://verfassungsblog.de/koennen-algorithmen-diskriminieren/; zu verschiedenen Bereichen, in denen Algorithmendiskriminierung problematisch ist: *Britz*: Einzelfallgerechtigkeit versus Generalisierung, S. 75 ff.
[112] *Orwat,* in *Antidiskriminierungsstelle des Bundes (Hrsg.)*: Diskriminierungsrisiken durch Verwendung von Algorithmen, S. 93; *Britz*: Einzelfallgerechtigkeit versus Generalisierung, S. 179 ff.

2.3.1 Schulische und universitäre Bildung

Im Sektor der schulischen und universitären Bildung gibt es bereits zahlreiche Bestrebungen, Lernprozesse zu standardisieren und personalisierte, auf den Förderbedarf der einzelnen Schüler bzw. Studierenden zugeschnittene Lernangebote zu schaffen.[113] Solche Programme lassen sich durchaus sinnvoll fruchtbar machen: So wäre es beispielsweise denkbar, mittels Big-Data-Analysen die Stärken, Schwächen und die Lerngeschwindigkeit der einzelnen Schüler und Studierenden zu erfassen, sie jeweils einem bestimmten Lernprofil zuzuordnen und ihnen auf dieser Basis ein individuelles Lernprogramm zuzuschneiden.[114]

Allerdings ebnet die für die Entwicklung personalisierter Lerninhalte erforderliche Standardisierung Ungleichbehandlungen den Weg. Gerade wenn die Lernprofile – der rasanten Entwicklung von Menschen im schulpflichtigen Alter wegen – nicht regelmäßig aktualisiert und unvoreingenommen angepasst werden, drohen Schüler Kategorien aufgezwängt zu bekommen, die ihrer Persönlichkeit und ihren individuellen Bedürfnissen nicht gerecht werden.

Ziel der Schulbildung ist, Menschen zu selbstbestimmtem und verantwortungsbewusstem Verhalten zu befähigen und ihnen reflexive Urteilskraft und Entscheidungsstärke beizubringen.[115] Bekommen die Schüler die Lerninhalte nur noch in individuell aufbereiteter Form zur Verfügung gestellt, könnte – je nach konkreter Ausgestaltung – die Fähigkeit des kritischen Hinterfragens von Informationen abhandenkommen. Wenn den Schülern durch algorithmisches Entscheiden eigene Entscheidungen abgenommen werden – man denke etwa an ein Programm, das präzise geeignete Ausbildungs- und Berufswege aufzeigen kann – droht die Herausbildung vernunftbegabter Wesen als zentrale Funktion schulischer Bildung nicht mehr gewährleistet zu sein.

Außerdem könnten Big-Data-Analysen auch dazu genutzt werden, Lehrkräften Empfehlungen für die Bewertung der Schüler und Studierenden zu geben. Das Verhalten der Lehrkräfte kann dabei durch den Automation Bias beeinflusst werden, wenn die Bewertungsvorschläge der Software nicht mehr kritisch hinterfragt werden.

2.3.2 Arbeitsmarkt

Auch auf dem Arbeitsmarkt sind Szenarien denkbar, in denen der Staat mithilfe von Big-Data-Analysen das Verhalten seiner Bürger steuert. So werden in Österreich[116] und Polen[117] bereits mittels prädiktiver Analysen die Arbeitsmarktchancen arbeitsloser Personen errechnet und kategorisiert. Zum einen können dadurch Diskriminierungen hervorgerufen werden.[118] Zum anderen kommt darüber hinaus vonseiten der Behörden aber auch hier wieder der Automation Bias zum Tragen, der bei automatisiertem Entscheiden stets problematisch ist.[119]

---

[113] Z.B. *Majumdar*, KI als helfende Lehrkraft während des Lockdowns, https://www.egovernment.de/ki-als-helfende-lehrkraft-waehrend-des-lockdowns-a-532bcc78c646262b120ca9bbf8771aaa/;
Deutscher Bildungsserver, Künstliche Intelligenz in der Schule,/. Entsprechende Anwendungen für Universitäten werden z.B. im BMBF-Projekt Komp-HI erforscht, https://www.bildungsserverkomp-hi.de/kuenstliche-intelligenz-in-der-schule-12990-de.html./.

[114] *Deutscher Ethikrat*: Mensch und Maschine, S. 166; vgl. *Damberger*, Lernende Schule, 20 (2017) 79, (22), 22; *Peters/Bovenschulte*: Learning Analytics, S. 1.

[115] *Deutscher Ethikrat*: Mensch und Maschine, S. 164.

[116] Zum österreichischen AMS *Fröhlich/Spiecker gen. Döhmann*, Können Algorithmen diskriminieren?, https://verfassungsblog.de/koennen-algorithmen-diskriminieren/.

[117] *Niklas/Sztandar-Sztanderska/Szymielewicz*, Profiling the Unemployed in Poland, https://panoptykon.org/biblio/profiling-unemployed-poland-social-and-political-implicationsalgorithmic-decison-making.

[118] Dazu *Fröhlich/Spiecker gen. Döhmann*, Können Algorithmen diskriminieren?, https://verfassungsblog.de/koennen-algorithmen-diskriminieren/; *Sesing/Tschech*, MMR 2022, 24 (24).

[119] Zum Teil wird daher vorgeschlagen, nicht von automatisiertem, sondern unterstütztem Entscheiden zu sprechen. Auf den Automation Bias hat der gewählte Begriff aber keine Auswirkungen, *Deutscher Ethikrat*: Mensch und Maschine, S. 226.

2.3.4 Gesundheitsbereich

Im Gesundheitsbereich stellen viele gesetzliche Krankenkassen bereits Programme für ihre Versicherten bereit, mit denen deren Gesundheitsdaten erfasst werden. So bieten viele Versicherer inzwischen Apps an, mit denen die Versicherten bequem Rechnungen, Rezepte oder ihre Krankmeldung an den Arbeitgeber einreichen können und über die sie – verknüpft mit dem Schrittzähler des Smartphones – von dem Versicherer mit Boni für ihre körperliche Aktivität „belohnt" werden können.[120] Inzwischen lässt sich mithilfe von Wearables wie Smartwatches der Gesundheitszustand einer Person relativ genau erfassen. Zudem wäre es denkbar, dass mithilfe des Smart Homes in Zukunft etwa auch anhand des Inhalts des Kühlschranks erfasst wird, wie einzelne Personen sich ernähren. Angesichts der bereits eingesetzten „Fitness-Tracker" ist durchaus denkbar, dass diese Daten von den gesetzlichen Krankenkassen[121] ausgewertet werden und auf dieser Grundlage durch Steuerungsmaßnahmen versucht wird, Menschen zu einem gesundheitsbewussteren Verhalten zu motivieren.

Dies scheint zunächst für beide Seiten nur von Vorteil zu sein: Der Staat einerseits könnte seine Ausgaben im Gesundheitssektor minimieren, während die Bürger ein gesünderes, längeres, vermeintlich glücklicheres Leben führen könnten. Problematisch ist dabei jedoch, wer definiert, was ein „gutes Leben" ausmacht. Überspitzt könnte man sagen: Wer bestimmt, dass ein Nichtraucher, der regelmäßig joggen geht und sich gesund ernährt, ein „besseres" Leben führt als ein Raucher, der seine Freizeit lieber auf dem Sofa verbringt und dabei exzessiv Schokolade isst?[122] Zweifel sind hier angebracht, auch wenn der Staat sich zu gesundheitsschädlichem Verhalten seiner Bürger nicht neutral verhalten muss[123] – im Gegenteil: Das BVerfG stuft die Volksgesundheit als „überragend wichtiges Gemeinschaftsgut" ein, das Eingriffe in entgegenstehende Grundrechte rechtfertigen kann.[124] Zudem obliegt dem Staat in Bezug auf die Volksgesundheit eine Schutzpflicht gegenüber seinen Bürgern, sie vor selbstschädigendem Verhalten zu bewahren.[125] Zumindest im Hinblick auf Selbstschädigungen, die unbestritten gesundheitsschädlich sind – wie Kettenrauchen oder übermäßiger Alkoholkonsum – ist es daher durchaus gerechtfertigt, dass der Staat ein Interesse hat, solches Verhalten zu unterbinden.[126]

Staatliche Steuerung im Gesundheitsbereich ist daher nicht per se unzulässig, sondern eine Frage der konkreten Ausgestaltung, der „Balance".[127] Staatliche Informationskampagnen für gesunde Ernährung oder Warnhinweise auf Tabakprodukten beispielsweise enthalten zwar eine eindeutige staatliche Wertung und dienen der Verhaltensbeeinflussung, lassen den Bürgern aber immer noch die Freiheit, dennoch zu rauchen und sich ungesund zu ernähren,[128] während ausbleibende Boni bei ungesundem Verhalten den Charakter einer handfesten Sanktion annehmen.

---

[120] Damit wirbt z.B. die Techniker Krankenkasse: *Die Techniker*, Belohnung durch Bewegung, https://www.tk.de/techniker/magazin/digitale-gesundheit/spezial/tk-fit-2066260?gclid=Cj0KCQjwslejBhDOARIsANYqkD0ucJ2AOY43pHy4O8HJ5_eWT7qm6vVKQj5U2nxP1xHQoBAf1R3AnbMaAuASEALw_wcB.

[121] Die gesetzlichen Krankenkassen sind als öffentlich-rechtliche Körperschaften durch Art. 1 III GG grundrechtsverpflichtet, vgl. *Wissenschaftliche Dienste Deutscher Bundestag*, Geltung der Grundrechte und des rechtsstaatlichen Rückwirkungsverbots für gesetzliche Krankenkassen, WD 3 – 3000 – 142/16, S. 4; *Schüffner/Franck* in *Sodan*, Handbuch des Krankenversicherungsrechts, § 36 Rdnr. 55.

[122] Vgl. *Wolff,* RW 2015, 194 (211).

[123] *Lübbe-Wolff* in *Kemmerer/Möllers/Steinbeis/Wagner (Hrsg.)*: Choice Architecture in Democracies, S. 248; vgl. *Kolbe*: Freiheitsschutz vor staatlicher Gesundheitssteuerung, S. 346 ff.

[124] BVerfGE 7, 377 (408) – „Apotheken-Urteil".

[125] BVerfGE 95, 173 (183 ff.) – „Warnhinweise auf Tabakerzeugnissen"; ablehnend bzgl. einer Pflicht zum Schutz des Einzelnen vor sich selbst *Kirste*, JZ 2011, 805 (813).

[126] So auch *Lübbe-Wolff* in *Kemmerer/Möllers/Steinbeis/Wagner (Hrsg.)*: Choice Architecture in Democracies, S. 249.

[127] *Lübbe-Wolff* in *Kemmerer/Möllers/Steinbeis/Wagner (Hrsg.)*: Choice Architecture in Democracies, S. 250; vgl. *Kolbe*: Freiheitsschutz vor staatlicher Gesundheitssteuerung, S. 348.

[128] Zwischen staatlichen Informationskampagnen und Warnhinweisen auf Tabakprodukten ist jedoch zu differenzieren: denn letztere verletzen zwar nicht die Freiheit der Bürger, verpflichten aber die Hersteller von Tabakprodukten, die Warnhinweise auf ihren Verpackungen abzudrucken und zwingen diese dadurch zu „nudgendem" Verhalten,

Zudem macht es einen Unterschied, wenn man sich vorstellt, dass solche staatlichen Steuerungsmaßnahmen auf Basis der umfassenden Auswertung personenbezogener Daten personalisiert werden. Dann könnten staatliche Informationskampagnen etwa individualisiert auf den Gesundheitszustand einzelner Bürger zugeschnitten werden. Raucher, die sich nach prädiktiven Analysen mit hoher Wahrscheinlich vom Nichtrauchen überzeugen lassen werden, würden mit anderen Informationen versorgt werden als Menschen, die sich eher nicht vom Rauchen abbringen lassen oder Nichtraucher. Wie bereits gezeigt, kann durch die Verarbeitung personenbezogener Daten grundsätzlich die Eingriffsintensität verstärkt werden.

2.3.5 Predictive Policing

Mithilfe von Predictive Policing lässt sich anhand von algorithmenbasierter Risikoanalysen die Wahrscheinlichkeit zukünftiger Straftaten vorhersagen, um etwa den Einsatz von Polizeikräften zu steuern oder die Rückfallrate verurteilter Straftäter zu bestimmen und Verbrechen zu verhindern. Während Predictive Policing in den USA bereits von einem Großteil der Sicherheitsbehörden angewendet wird, kommt es in Deutschland bisher nur vereinzelt zum Einsatz, war jedoch in den letzten Jahren zunehmend Thema im politischen und wissenschaftlichen Diskurs. Bayern und Baden-Württemberg beispielsweise setzen die Software Pre Crime Observation System (PRECOBS) ein, um Wohngebiete zu ermitteln, in denen eine erhöhte Wahrscheinlichkeit für Einbruchdiebstähle besteht.[129] In den USA hingegen wird Predictive Policing bereits nicht mehr nur raum- sondern auch personenbezogen eingesetzt, um für einzelne Personen Risikoprofile zu erstellen.[130]

Zum einen birgt Predictive Policing die Gefahr, dass sich der menschliche Entscheider, der die Ergebnisse der Software auswertet, vorbehaltlos den maschinellen Empfehlungen anschließt (Automation Bias). Zudem droht Predictive Policing Verzerrungen in Bezug auf strafbares Verhalten zu reproduzieren und zu perpetuieren. Es ist allgemein bekannt, dass die Strafverfolgungsbehörden nur einen Bruchteil des strafbaren Verhaltens erfassen und dadurch bestimmte Straftaten eher wahrgenommen werden als andere.[131] Diese Verzerrungen können durch Predictive Policing reproduziert und dadurch verstärkt werden. Während einzelne Polizeibeamte sich ihrer eigenen Subjektivität in der Regel bewusst sind, erwecken Algorithmen den Anschein, besonders objektive und fundierte Ergebnisse zu liefern.[132] Algorithmische Entscheidungen sind jedoch keinesfalls fehlerfrei. Abgesehen von den einhergehenden Diskriminierungen tragen zudem auch falsch-positive oder falsch-negative Treffer dazu bei, dass Verzerrungen perpetuiert werden.[133] Zwar machen auch menschliche Entscheider Fehler; durch die Schaffung neuer Querverbindungen in großen Datensätzen können aber gänzlich neue Verdachtsmomente entstehen, die sich dann breitenwirksam entfalten.[134] Besonders folgenschwer ist es, wenn aufgrund falscher Treffer (operative) Folgemaßnahmen gegen den Einzelnen ergriffen werden, die ex post nicht gerechtfertigt sind.[135]

---

s. dazu *Lübbe-Wolff* in *Kemmerer/Möllers/Steinbeis/Wagner (Hrsg.)*: Choice Architecture in Democracies, S. 250; dazu, dass der Staat vorrangig aufklären und warnen sollte *Kolbe*: Freiheitsschutz vor staatlicher Gesundheitssteuerung, S. 348 f; Warnen und Aufklären seien keine paternalistischen Handlungen, *Kirste*, JZ 2011, 805 (813).

[129] *Singelnstein,* NStZ 2018, 1 (1); *Seidensticker/Bode/Stoffel*: Predictive Policing in Germany, S. 3; in Bezug auf Baden-Württemberg *Haake*, WISTA 2/2021, 59 (59 f.).

[130] *Singelnstein* NStZ 2018, 1 (2); *Ferguson*, University of Pennsylvania Law Review 2015, Vol. 163, 327 (373); *Eisele/Böhm*, in *Beck/Kusche/Valerius (Hrsg.):* Digitalisierung, Automatisierung, KI und Recht, S. 526.

[131] *Singelnstein,* NStZ 2018, 1 (4).

[132] *Singelnstein,* NStZ 2018, 1 (4); *Deutscher Ethikrat*: Mensch und Maschine, S. 241 ff; *Kuhlmann/Trute*, GSZ 2021, 103 (106).

[133] *Deutscher Ethikrat*: Mensch und Maschine, S. 244; vgl. *Rademacher/Perkowski*, JuS 2020, 713 (716); *Kuhlmann/Trute*, GSZ 2021, 103 (110).

[134] Vgl. insofern auch BVerfGE 156, 11 (40) – „Antiterrordateigesetz II"; *Gutjahr/Limberger,* DÖV 2022, 848 (854).

[135] *Gutjahr/Limberger,* DÖV 2022, 848 (854); vgl. *Rademacher/Perkowski*, JuS 2020, 713 (716).

Nicht zuletzt kommen im Bereich des Predictive Policing aufgrund der drohenden Strafverfolgung oder operativer Eingriffe Verhaltensbeeinflussungen durch Chilling Effects in besonderem Maße zum Tragen.[136]

2.4 Fazit

Staatliche Verhaltensbeeinflussungen sind also nicht prinzipiell verboten, sie müssen aber zumindest die Entscheidungsfreiheit des Einzelnen aufrechterhalten. Dafür können Big-Data-Analysen zweifellos ein sinnvolles Hilfsmittel darstellen. Die Technologie bietet hier große Potenziale. Allerdings verstärkt die Auswertung von Big Data regelmäßig die Intensität von Grundrechtseingriffen. Eine Kombination von staatlicher Steuerung und Big-Data-Auswertungen ist daher im Hinblick auf grundrechtliche Gewährleistungen besonders problematisch.

[136] Zur Verhaltensbeeinflussung durch Beobachtung *Büscher et al.*: Verhaltensbeeinflussung durch Beobachtung?; *Schlehahn/Aichroth/Mann/Schreiner/Shepherd/Wong*: Benefits and Pitfalls of Predictive Policing, S. 2.

3. KI und große, reale Datenmengen

3.1 Herausforderungen großer Datenmengen

Die Verarbeitung großer, realer Datenmengen ist ein ideales Einsatzfeld für die sogenannte „künstliche Intelligenz" (im Folgenden: „KI"). Zum von der EU-Kommission vorgelegten Entwurf der KI-Verordnung der Europäischen Union[137] (im Folgenden: „KI-VO-E"[138]) liegt inzwischen auch ein Änderungsvorschlag des EU-Parlaments vor. KI wird in Art. 3 Nr. 1 KI-VO-E definiert als „eine Software, die mit einer oder mehreren der in Anhang I aufgeführten Techniken und Konzepten[139] entwickelt worden ist und im Hinblick auf eine Reihe von Zielen, die vom Menschen festgelegt werden, Ergebnisse wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das Umfeld beeinflussen, mit dem sie interagieren." Erwgr. 3 des KI-VO-E beschreibt KI als „eine Reihe von Technologien, die sich rasant entwickeln und zu vielfältigem Nutzen für Wirtschaft und Gesellschaft über das gesamte Spektrum industrieller und gesellschaftlicher Aktivitäten hinweg beitragen können. Durch die Verbesserung der Vorhersage, Optimierung der Abläufe, Ressourcenzuweisung und Personalisierung digitaler Lösungen, die Einzelpersonen und Organisationen zur Verfügung stehen, kann die Verwendung künstlicher Intelligenz den Unternehmen wesentliche Wettbewerbsvorteile verschaffen und zu guten Ergebnissen für Gesellschaft und Umwelt führen, beispielsweise in den Bereichen Gesundheitsversorgung, Landwirtschaft, allgemeine und berufliche Bildung, Infrastrukturmanagement, Energie, Verkehr und Logistik, öffentliche Dienstleistungen, Sicherheit, Justiz, Ressourcen- und Energieeffizienz sowie Klimaschutz und Anpassung an den Klimawandel."

Um diese „guten Ergebnisse" zu erzielen, benötigt die KI einerseits große Datenmengen als Trainingsdatenpool zur Analyse von Mustern und logischen Zusammenhängen, zum anderen ist die anschließende Verarbeitung großer Datenmengen durch KI deren ideales Anwendungsfeld. Die Existenz großer Datenmengen zur Verarbeitung kann verschiedene rechtliche Implikationen nach sich ziehen: Neben in diesem Kapitel nicht behandelten datenschutzrechtlichen Herausforderungen[140] stellen sich Fragen aus dem Bereich der Validität der Daten, der Diskriminierungsfreiheit, der Rechte an den Daten und der haftungsrechtlichen Konsequenzen. Erwgr. 2a des KI-VO-E in der Version des Europäischen Parlaments[141] weist speziell auf Risiken großer Datenmengen für die Persönlichkeitsrechte hin: „Da künstliche Intelligenz oft auf die Verarbeitung großer Datenmengen angewiesen ist und viele KI-Systeme und -Anwendungen auf der Verarbeitung personenbezogener Daten beruhen, sollte sich diese Verordnung auch auf Artikel 16 AEUV stützen, in dem das Recht auf den Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten verankert ist und der den Erlass von Vorschriften zum Schutz von Privatpersonen im Hinblick auf die Verarbeitung personenbezogener Daten vorsieht." Je größer die Datenmengen sind, umso weniger spielen statistische Ausreißer eine Rolle, da der Prozentsatz der Ausreißer entsprechend sinkt, soweit die gesamte Datenmenge betrachtet wird.[142] Zugleich können diese bei größeren Datenmengen umso wahrscheinlicher Bestandteil des

[137] Vorschlag für eine Richtlinie des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften über künstliche Intelligenz (Gesetz über Künstliche Intelligenz), COM/2021/206 final.
[138] Soweit im Folgenden von Artikeln des KI-VO-Entwurfs die Rede ist, handelt es sich um Regelungen, die im Kommissionsentwurf und im Entwurf des EU-Parlaments identisch geregelt sind.
[139] Anhang 1 benennt unter anderem Konzepte des maschinellen Lernens, mit beaufsichtigtem, unbeaufsichtigtem und bestärkendem Lernen unter Verwendung einer breiten Palette von Methoden, einschließlich des tiefen Lernens (Deep Learning) sowie Logik- und wissensgestützte Konzepte und Statistische Ansätze, Bayessche Schätz-, Such- und Optimierungsmethoden.
[140] Siehe dazu unten, S. 39 ff.
[141] https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_DE.html.
[142] Dies gilt jedenfalls, wenn das Anwachsen der Datenmenge repräsentativ für die potentielle Gesamtmenge ist.

Datenpools sein. Falls diese Ausreißer als einzelne Datensätze rechtlich problembehaftet sind, sind Lösungen für den Umgang mit diesen Daten zu finden, etwa durch Anwendungen, welche in der Lage sind, entsprechende einzelne Datensätze auszufiltern. Beispiele für eine Problembehaftung könnten einzelne Datensätze sein, welche urheberrechtlichen Unterlassungsansprüchen ausgesetzt sind, Persönlichkeitsrechte verletzen oder aufgrund ihrer sachlichen Fehlerhaftigkeit bei der Verarbeitung zu unrichtigen Ergebnissen führen und somit haftungsrechtlich problematisch sind. Zudem könnten unerwünschte Positionen verbreitet werden, wenn sie – etwa aufgrund von Troll-Aktivitäten und Fake-Accounts / Chatbots – in großem Umfang in die Trainingsdaten einfließen. Abhängig vom Einsatzzweck der KI kann die Verarbeitung großer Datenmengen dazu führen, dass sich regulatorische Zuordnungen nach der künftigen KI-Verordnung verändern.

Im Folgenden soll aufgezeigt werden, welche rechtlichen Herausforderungen sich bei der Gewinnung und der Verarbeitung großer Datenmengen stellen, welche Herausforderungen sich ergeben und welche Neuregelungen der EU zu erwarten sind, die bestimmenden Einfluss auf entsprechende KI-Systeme haben werden.

3.2 Rechtmäßigkeit der Datengewinnung aus Sicht der geistigen und gewerblichen Schutzrechte

Bei großen Datenmengen stellt sich die Frage, wie diese in rechtmäßiger Weise nutzbar gemacht werden können, soweit sie nicht aus Verträgen mit den Inhabern von Datenpools (etwa Bild- oder Textdatenbanken) stammen, die ihrerseits zur Weitergabe berechtigt sind. Die sich für große Mengen vorrangig anbietende Möglichkeit der Datengewinnung ist das Auslesen und Kopieren von Daten aus nicht zugangsbeschränkten Internetseiten (sog. „Screenscraping“ oder „Webscraping“). Diese ohne Einwilligung der Seiteninhaber durchgeführte „Datenabschöpfung” kann jedoch eine Verletzung von Schutzrechten geistigen oder gewerblichen Eigentums darstellen. Für die Nutzung der Daten ist daher von großer Bedeutung, dass diese nicht Rechte Dritter verletzt und die Auswertung der großen Datenmengen nicht entsprechenden Unterlassungsansprüchen ausgesetzt ist[143]. Ausgelesene Daten könnten Werke nach § 2 UrhG in Form etwa von Sprachwerken, Musik, Lichtbildern oder Darstellungen wissenschaftlicher oder technischer Art enthalten, deren Vervielfältigung und Verbreitung exklusiv dem Urheber nach § 15 Abs.1 UrhG vorbehalten ist.

In Betracht kämen weiterhin urheberrechtliche Unterlassungsansprüche der Webseitenbetreiber nach § 97 Abs 1 UrhG. Soweit die fremden ausgelesenen Webseiten Datenbanken nach § 87a UrhG enthalten, also eine „Sammlung von Werken, Daten oder anderen unabhängigen Elementen, die systematisch oder methodisch angeordnet und einzeln mit Hilfe elektronischer Mittel oder auf andere Weise zugänglich sind und deren Beschaffung, Überprüfung oder Darstellung eine nach Art oder Umfang wesentliche Investition erfordert” (§ 87a Abs. 1 S. 1 UrhG), unterliegt die Übernahme der Datenbank oder Teilen derselben den Vorgaben des § 87b UrhG. Nach dessen Satz 2 kann das systematische wiederholte Auslesen einen Verstoß darstellen, „sofern diese Handlungen einer normalen Auswertung der Datenbank zuwiderlaufen oder die berechtigten Interessen des Datenbankherstellers unzumutbar beeinträchtigen”. Eine solche unzumutbare Beeinträchtigung ist zu bejahen, wenn „die Entnahmehandlungen hierauf gerichtet sind und im Fall ihrer Fortsetzung dazu führen würden, die Datenbank insgesamt oder einen nach Art oder Umfang wesentlichen Teil zu vervielfältigen, zu verbreiten oder öffentlich wiederzugeben und damit den Tatbestand des Art. 7 Abs. 1 der Datenbank-RL zu umgehen“.[144] Je nach Systematik und Umfang der Übernahme von Daten aus einer gegen den sogenannten

[143] *Käde*, Kreative Maschinen und Urheberrecht, S. 72 zum Einwilligungserfordernis.
[144] BGH, 1. 12. 2010 – I ZR 196/08, K&R 2011, 485 ff.

„Robot"-Zugriff technisch nicht geschützten Seite[145] ist daher ein Verstoß gegen das Leistungsschutzrecht der Datenbanken möglich. Erfüllen die ausgelesenen Daten ausnahmsweise aufgrund besonders kreativer Auswahl der Einzelelemente die Vorgaben eines Datenbankwerks nach § 4 UrhG, stellt das ungenehmigte Vervielfältigen einen Verstoß gegen das Ausschließlichkeitsrecht des Urhebers nach § 15 Abs. 1 Nr. 1 UrhG dar. Allerdings dürfen auch urheberrechtlich geschützte Elemente als Trainingsdaten übernommen werden, soweit die Anwendung des in Umsetzung der DSM-Richtlinie[146] eingeführten § 44b UrhG zum Text- und Data Mining in Betracht kommt. Nach dessen Satz 1 ist die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken erlaubt, um daraus Informationen insbesondere über „Muster, Trends und Korrelationen zu gewinnen". Nach § 44b Abs. 2 UrhG sind Vervielfältigungen von rechtmäßig zugänglichen Werken für das Text- und Data Mining zulässig. Die Vervielfältigungen sind zu löschen, wenn sie für das Text- und Data Mining nicht mehr erforderlich sind. Nach Abs. 3 des § 44b UrhG sind solche Nutzungen nach Abs. 2 S. 1 nur zulässig, wenn der Rechtsinhaber sich diese nicht vorbehalten hat. Ein Nutzungsvorbehalt bei online zugänglichen Werken ist nur dann wirksam, wenn er in maschinenlesbarer Form erfolgt.

Hat der Webseitenbetreiber – soweit er Schutzrechtsinhaber ist – keinen maschinenlesbaren Vorbehalt auf seiner Seite vorgesehen (also „jede digital hinterlegte Information als Vorbehalt […], die in einem Internetstandard für Text codiert ist"[147]), kann sich der Betreiber der KI für den Vorgang des Auslesens und der Analyse auf die Schranke des § 44b UrhG berufen. Wird allerdings ein solcher Vorbehalt auf der Webseite im Impressum oder im Seitenquelltext in den Metatags „Robot" und „Copyright" hinterlegt, muss der KI-Betreiber dies erkennen und auf das Auslesen der Daten zum Text- und Data Mining verzichten.

Ob unabhängig von urheberrechtlichen Vorgaben durch die Geltendmachung eines „virtuellen Hausrechts" (ohne technische Abwehr) das Auslesen von Webseiten auch AGB-rechtlich wirksam untersagt werden kann, ist umstritten[148], aus dem Besuch einer Webseite alleine lässt sich im Regelfall noch kein Wille zum Abschluss eines Vertrags ableiten.[149] Ein Verstoß gegen wirksam eingebundene Nutzungsbedingungen wird in aller Regel auch nicht als unlautere Behinderung gem. § 4 Nr. 10 UWG zu werten sein[150]. Die technische Verhinderung des Auslesens stellt keinen Verstoß gegen § 44 b UrhG dar, auch wenn kein Vorbehalt im Sinne des § 44b Abs. 3 UrhG erklärt wurde[151]. Wird bei einem Seitenbesuch durch Robots eine technische Abwehrmaßnahme umgangen, kann dies allerdings einen Wettbewerbsverstoß darstellen[152]. Es ist nicht erforderlich, dass die ausgelesenen Werke veröffentlicht worden sind, es genügt, dass sie für die Auslesenden technisch zugänglich waren[153]. Die Digitalisierung der betroffenen Werke kann auch erst von demjenigen vorgenommen worden sein, der sich auf die Schranke des § 44b UrhG beruft[154].

3.3 Rechtsfragen der Verarbeitung großer Daten

Sind die Daten rechtmäßig gewonnen worden, ist zu prüfen, ob sie auch rechtmäßig verarbeitet werden. Dies beinhaltet die Prüfung, ob eine Verletzung von Rechten Dritter in

[145] Wobei in der Tatsache, dass ein Webseitenbetreiber seine Seite nicht gegen Robot-Zugriff geschützt hat, wohl keine Einwilligung in das Auslesen von Daten im Sinne einer Nutzungsrechtsübertragung zu sehen sein wird.
[146] Richtlinie über das Urheberrecht und die verwandten Schutzrechte im digitalen Binnenmarkt und zur Änderung der Richtlinien 96/9/EG und 2001/29/EG.
[147] *Jacobsen/Hartmann*, MMR-Aktuell 2021, 441332.
[148] BGH, 30. 4. 2014 – I ZR 224/12, K&R 2014, 596, Rn. 35, 38 – Screen Scraping; *Kianfar*, DSRITB 2014, 821.
[149] BGH, 30. 4. 2014 – I ZR 224/12, K&R 2014, 596, Rn. 37 – Screen Scraping.
[150] BGH, 30. 4. 2014 – I ZR 224/12, K&R 2014, 596, Rn. 35, 38 – Screen Scraping.
[151] *Dreier* in *Dreier/Schulze (Hrsg.)*, Urheberrechtsgesetz, § 44b UrhG, Rdnr. 10.
[152] BGH, 30. 4. 2014 – I ZR 224/12, K&R 2014, 596, Rn. 37 – Screen Scraping.
[153] *Dreier* in *Dreier/Schulze (Hrsg.)*, Urheberrechtsgesetz, 7. Auflage 2022, § 44b UrhG, Rdnr. 6.
[154] *Bullinger* in *Wandtke/Bullinger (Hrsg.)*, Urheberrecht, 6.A. 2022, § 44b UrhG, Rdnr. 4.

Betracht kommt und welche eigenen Rechte KI-Anbietern an den KI-Ergebnissen zustehen können.

3.3.1 Schrankenregelungen der §§ 44b, 60d UrhG

Im Rahmen der Verarbeitung urheberrechtlich geschützter Datensätze stellt sich die Frage der Reichweite der Schrankenregelungen zum Text- und Data Mining der §§ 44b, 60d UrhG.

Im Rahmen der Anwendung künstlicher Intelligenz werden zur Auswertung des Trainingsdatenpools Methoden eingesetzt, welche auf der Basis von sog. „Clusterings"[155] die neu geordnete Produktion der KI-Ergebnisse ermöglichen[156]. Nach § 44b Abs. 1 UrhG ist – entsprechend der Definition in Art. 2 Nr. 2 der DSM-Richtlinie – Text- und Data Mining „die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen".

Die Auswertung der erhobenen Daten, die Annotation zu bestimmten Begriffen und die entsprechende Neustrukturierung kann als Teil des Text- und Data Minings im Sinne des § 44b UrhG verstanden werden[157]. Greift die Schranke des § 44b UrhG, wird das Bearbeitungsrecht des Urhebers nach § 23 UrhG nicht verletzt[158]. Eine Pflicht zur Nennung der Urheber der ausgelesenen Texte besteht nicht[159], ebenso wenig ist eine Vergütung erforderlich[160]. Ausgewertet werden können Werke und verwandte Schutzrechte, somit unter anderem auch wissenschaftliche Ausgaben, Lichtbilder, Darbietungen ausübender Künstler, Tonträger Sendungen, Datenbanken, Presseveröffentlichungen und Laufbilder. Nach § 44b Abs. 2 S. 2 UrhG sind die Vervielfältigungen allerdings zu löschen, wenn sie für das Text- und Data Mining nicht mehr erforderlich sind. Das Text- und Data Mining erfordert jedoch eine ständige Verfügbarkeit des Trainingsdatenpools, sodass vertreten wird, dass die Erforderlichkeit während der Laufzeit des Betriebs des Minings nicht entfällt[161]; mithin wäre eine Löschung erst erforderlich, wenn der KI-Betrieb mit diesem Datenpool eingestellt wird.

Neben der Schranke des § 44b UrhG kommt für KI-Auswertungen für Forschungszwecke auch die Schranke des § 60d UrhG in Betracht. Nach diesem sind solche – nach § 44b Abs. 1 UrhG definierten – Vervielfältigungen für Text- und Data Mining für Zwecke der wissenschaftlichen Forschung zulässig, „wenn Forschungsinstitutionen, sofern sie nicht kommerzielle Zwecke verfolgen, sämtliche Gewinne in die wissenschaftliche Forschung reinvestieren oder im Rahmen eines staatlich anerkannten Auftrags im öffentlichen Interesse tätig sind". Ausgeschlossen sind nach § 60d Abs. 2 S. 2 UrhG Forschungsorganisationen, die mit einem privaten Unternehmen zusammenarbeiten, das einen bestimmenden Einfluss auf die Forschungsorganisation und einen bevorzugten Zugang zu den Ergebnissen der wissenschaftlichen Forschung hat. Nach § 60d Abs. 3 Nr. 2 UrhG sind auch einzelne Forscher berechtigt, sofern sie nicht kommerzielle Zwecke verfolgen.

Ein Ausschluss des § 60d UrhG durch einen Nutzungsvorbehalt entsprechend § 44b UrhG ist nicht vorgesehen, sodass die KI-Bearbeitung großer Datenmengen hier uneingeschränkt – allerdings beschränkt auf Forschungszwecke – möglich ist. Die Berechtigten

[155] *Schildt* in *Kaulartz/Braegelmann (Hrsg.)*, Rechtshandbuch Artificial Intelligence und Machine Learning, Kap. 15.3, Rn. 15.
[156] Siehe zum Analyseverfahren auch *Dehio/Reul* in *Chibanguza/Kuß/Steege (Hrsg.)*, Künstliche Intelligenz, Kap. A, R. 33f.
[157] *Dreier* in *Dreier/Schulze (Hrsg.)*, Urheberrechtsgesetz, § 44b UrhG, Rn. 5.
[158] *Bullinger* in *Wandtke/Bullinger (Hrsg.)*, Urheberrecht, § 44b UrhG, Rn. 5.
[159] *Dreier* in *Dreier/Schulze*, Urheberrechtsgesetz, § 44b UrhG, Rn. 10.
[160] *Bullinger* in *Wandtke/Bullinger (Hrsg.)*, Urheberrecht, § 44b UrhG, Rn. 11; *Dreier* in *Dreier/Schulze (Hrsg.)*, Urheberrechtsgesetz, § 44b UrhG, Rn. 14.
[161] *Bullinger* in *Wandtke/Bullinger (Hrsg.)*, Urheberrecht, § 44b UrhG, Rn. 9.

dürfen nach § 60d Abs. 5 UrhG Vervielfältigungen mit angemessenen Sicherheitsvorkehrungen gegen unbefugte Benutzung aufbewahren, solange sie für Zwecke der wissenschaftlichen Forschung oder zur Überprüfung wissenschaftlicher Erkenntnisse erforderlich sind. Die aus einem entsprechend nach § 60d UrhG berechtigten Data Mining gewonnenen Vervielfältigungen dürfen einem bestimmt abgegrenzten Kreis von Personen für deren gemeinsame wissenschaftliche Forschung sowie einzelnen Dritten zur Überprüfung der Qualität wissenschaftlicher Forschung nach § 60d Abs. 4 S. 1 UrhG öffentlich zugänglich gemacht werden im Sinne des § 19a UrhG. Anders als bei den Berechtigten nach § 44b UrhG ist eine kommerzielle Nutzung nicht zulässig[162]. Diese öffentliche Zugänglichmachung ist entsprechend § 60d Abs, 4 S. 2 UrhG zu beenden, sobald die gemeinsame wissenschaftliche Forschung oder die Überprüfung der Qualität wissenschaftlicher Forschung abgeschlossen ist.

3.3.2 Verletzung von Rechten Dritter

Soweit das Auslesen der Daten rechtmäßig erfolgt ist, stellt sich dennoch die Frage, ob eine Verletzung von Rechten Dritter durch die Verarbeitung innerhalb der KI und Ergebnisdarstellung durch die KI in Betracht kommt. Ergebnisse der KI könnten die Rechte an fremden Werken (außerhalb des Datenbankschutzes) verletzen, wenn diese Werke nur kaum verändert vervielfältigt würden und eine unfreie Bearbeitung nach § 23 UrhG gegeben wäre. KI-Ergebnisse könnten fremde Rechte verletzen, wenn diese umfangreiche Anteile von Vorarbeiten Dritter enthielten[163]. Diese müssten unmittelbar in generierte Bilder, Software, Musik, Sprache oder Text einfließen, was am ehesten bei Bildern denkbar ist[164], wenn bei Grafikerstellungen konkrete von Menschen vorgefertigte Bestandteile übernommen und integriert werden. Meist beruhen KI-Modelle auf Algorithmen, die in der Lage sind, sowohl große Datenmengen auszuwerten als auch sie inhaltlich zu bewerten und fortlaufend veränderliche Ergebnisse wiederzugeben, wobei der Grad der Autonomie – insbesondere bei neuronalen Netzwerken[165] – das besondere Charakteristikum ist[166]. Je größer die Autonomie, umso weniger Übernahmen vorgefertigter Bestandteile werden zu finden sein, die einem Urheberrechtsschutz zugänglich wären. Grundsätzlich ist daher eine Wiedergabe vorgefertigter Werkanteile Dritter unwahrscheinlich. Ausnahmen kommen in Betracht, wenn die KI es zulässt, dass als KI-Ergebnis eine Reproduktion etwa eines Bildes aus dem Trainingsdatenpool erstellt wird, welche dem Original so nahekommt, dass sie als unfreie Bearbeitung nach § 23 Abs. 1 UrhG zu werten wäre[167].

Außerhalb des Urheberrechts können Daten verschiedenen weiteren Regulierungen unterfallen. Soweit Daten durch den KI-Hersteller aus öffentlich zugänglichen Quellen bezogen werden, kommt aufgrund des dann offensichtlich fehlenden Schutzes der Daten vor den Zugriffen Dritter kein Anspruch der Dateninhaber aus dem GeschGehG in Betracht, da die Daten in diesem Fall nach § 2 Abs. 1 b) GeschGehG nicht „Gegenstand von den Umständen nach angemessenen Geheimhaltungsmaßnahmen durch ihren rechtmäßigen Inhaber" waren. Auch Ansprüche aus Eigentum nach § 929 BGB oder aus Besitzrechten nach § 854 BGB kommen bei reinen Datensätzen nicht in Betracht[168]. Markenrechtliche Ansprüche Dritter können der Verarbeitung großer Datensätze nur dann entgegenstehen, wenn die KI vorhandene Bild- oder Wortbildmarken wenig verändert wiedergibt, sodass markenrechtliche Ansprüche nach § 14 Abs. 2 Nr. 2 MarkenG (Verwechslungsschutz) oder § 14 Abs. 2 Nr. 3 MarkenG (Bekanntheitsschutz) in Betracht kommen. Dies setzt jedoch voraus, dass im Fall des § 14 Abs. 2 Nr. 2 MarkenG die Wiedergabe der Marke in einer ähnlichen Waren- oder Dienstleistungsklasse erfolgt bzw. es sich im Fall des § 14

[162] *Ulmer-Eilfort/Druschel* in *Ulmer-Eilfort/Obergfel (Hrsg.)*, Verlagsrecht, Kap. G, Rn. 820.
[163] *Loewenheim/Leistner* in *Schricker/Loewenheim (Hrsg.)*, Urheberrecht, § 2 UrhG Rn. 39.
[164] Beispiele bei *Wilmer*, K&R 2023, 385 (387).
[165] Siehe zur Definition *Söbbing*, MMR 2021, 111.
[166] *Loewenheim/Leistner* in *Schricker/Loewenheim (Hrsg.), Urheberrecht*, § 2 UrhG Rn. 41, 41a.
[167] Beispiele bei *Wilmer*, K&R 2023, 385 (387).
[168] *Kühling/Sackmann*, ZD 2020, 24 (26 f).

Abs. 2 Nr. 3 MarkenG bei der Marke um eine im Inland bekannte Marke handelt und die Benutzung des Zeichens die Unterscheidungskraft oder die Wertschätzung der bekannten Marke ohne rechtfertigenden Grund in unlauterer Weise ausnutzt oder beeinträchtigt. Denkbar wäre in seltenen Fällen auch, dass ein Verfall der Marke nach § 49 Abs. 2 Nr. 1 MarkenG durch eine Verstärkung der Entwicklung des Markenbegriffs zur gebräuchlichen Bezeichnung für das Produkt oder die Dienstleistung begünstigt wird. Daher wird bei der Verarbeitung großer Datenmengen die Klärung der Urheberrechte Dritter neben patentrechtlichen Themen (vor allem in den USA) im Vordergrund stehen.

3.3.3 Schutz der KI-Ergebnisse zugunsten des KI-Betreibers

Soweit KI-Anbieter in die Verarbeitung großer Datenmengen investieren, stellt sich die Frage, inwiefern sie – nach Absicherung der Zulässigkeit der Gewinnung der Trainingsdaten und der Freiheit der Ergebnisse von Rechten Dritter – selbst Rechte an den Ergebnissen postulieren können. Ein urheberrechtlicher Schutz eines Werks setzt nach § 2 Abs. 2 UrhG eine persönlich-geistige Schöpfung voraus, welche auf einer menschlichen Handlung beruht[169]. Gleiches gilt für den Schutz als Sammelwerk nach § 4 Abs. 1 UrhG oder als Datenbankwerk nach § 4 Abs. 2 UrhG; auch hier ist menschliches Handeln erforderlich[170]. In Ausnahmefällen wäre dies denkbar, soweit im Rahmen der „Promptografie" eine stark in Algorithmen oder Datensätze eingreifende menschliche Gestaltung vorliegt, welche sich der KI als Werkzeug bedient.[171] Denkbar wäre auch ein Datenbankschutz nach § 87a ff. UrhG für die hinter dem KI-Modell stehenden Teile der Modellierung, da als Datenbankhersteller – anders als beim Werk – auch eine juristische Person in Betracht käme. Da eine Datenbank nach § 87a Abs. 1 S. 1 UrhG eine Sammlung von Werken, Daten oder anderen unabhängigen Elementen ist, die systematisch oder methodisch angeordnet und einzeln mithilfe elektronischer Mittel oder auf andere Weise zugänglich sind und deren Beschaffung, Überprüfung oder Darstellung eine nach Art oder Umfang wesentliche Investition erfordert, könnten zwar Trainingsdaten als Datenbank eingeordnet werden, allerdings nicht die durch die KI erzeugten Ergebnisse[172].

Teilweise wird de lege ferenda ein eigenes Schutzrecht für KI-generierte Inhalte gefordert, das jedoch auf den Schutz der Trainingsdaten beschränkt sein soll[173] und damit nicht für den Schutz der KI-Ergebnisse anwendbar wäre. Deren Schutz könnte jedoch grundsätzlich vertraglich de lege lata abgesichert werden, sodass es im Ermessen der Betreiber steht, die Ergebnisse wirtschaftlich zu verwerten. Soweit allerdings vertraglich Rechte an nicht schutzfähigen Ergebnissen eingeräumt werden, stellt sich die Frage, ob dies nach deutschem AGB-Recht zulässig ist, da eine Lizenzierung nicht geschützter Inhalte einen Verstoß gegen § 307 Abs. 2 Nr. 1 BGB darstellen könnte. Denkbar wäre eine Lizenzierung der Nutzung der Software anstelle der KI-Ergebnisse als solchen, etwa im Rahmen eines Pachtvertrags nach § 581 Abs. 2 BGB[174] oder eines typengemischten Vertrags mit kauf- und dienstvertraglichen Elementen, gegebenenfalls auch eines Vertrags sui generis zur KI-Nutzung[175]. Patentrechtliche Schutzmöglichkeiten setzen einen menschlichen Erfinder voraus[176] und werden darüber hinaus an den Anforderungen an die Technizität der Erfindung scheitern[177].

Verbleibende Schutzmöglichkeiten sind somit für die KI-Software als solche die §§ 69a UrhG, für KI-Trainingsdaten unter bestimmten Umständen §§ 2 Abs. 1 Nr. 7, § 4 Abs. 2

[169] *Bullinger* in: *Wandtke/Bullinger (Hrsg.)*, UrhG, UrhG, § 2 Rn. 15; *Loewenheim/Leistner* in *Schricker/Loewenheim (Hrsg.)*, Urheberrecht, § 2 UrhG Rn. 38 m. w. N.
[170] *Käde*, Kreative Maschinen und Urheberrecht, S. 176 ff.
[171] Einzelheiten und Fallbeispiele mit weiteren Nachweisen bei Wilmer, S. 389 f.
[172] Vgl. *Vogel* in *Schricker/Loewenheim (Hrsg.),* Urheberrecht, § 87a UrhG Rn. 34.
[173] *Hacker*, GRUR 2020, 1025 (1033).
[174] *Hennemann*, RDi 2021, 61 (64) m. w. N.
[175] *Hennemann*, RDi 2021, 61 (64).
[176] *Hetmank/Lauber-Rönsberg*, GRUR 2018, 574 (575).
[177] *Hetmank/Lauber-Rönsberg*, GRUR 2018, 574 (576).

und § 87a UrhG[178], für KI-Modelle das GeschGehG[179] oder der ergänzende wettbewerbliche Leistungsschutz nach § 4 Abs. 3 UWG[180].

3.4 Haftungsrechtliche Fragen de lege lata

Bei großen Datenmengen können sich für die Haftung für die KI-Ergebnisse besondere Herausforderungen stellen. Einerseits kann eine große Datenmenge die Zuverlässigkeit der Datenbasis erhöhen, indem statistische Ausreißer in der großen Datenmenge eine geringere Bedeutung erhalten. Andererseits erhöht sich die Wahrscheinlichkeit, dass sich unter den zahlreichen Datensätzen auch solche finden, welche als Einzeldatensatz rechtlich problematisch sind. Dies kann sich sowohl auf Rechtsverletzungen durch Bildinhalte als auch auf problematische Textinhalte beziehen, welche etwa sog. „Hate Speech" enthalten. Hinzu kommt das Problem der Variabilität der Datensätze, soweit eine KI nicht auf einer statischen Datenmenge beruht, die nur in größeren Abständen aktualisiert wird. Die Haftungsfragen können sowohl KI-Betreiber betreffen als auch die Lieferanten von Datenbeständen und/oder einzelnen Datensätzen. Schäden könnten etwa in der auf falschen Daten beruhenden fehlerhaften Ausgabe von Handlungsempfehlungen im Bereich der Medizin oder bei Anwendungen im Bereich des Börsenhandels entstehen, aber auch in zahlreichen anderen Einsatzbereichen. Dies gilt vor allem für eine generalisierte KI, welche über Schnittstellen in alle denkbaren Anwendungen integriert werden kann[181]. Gegenüber Nutzern, mit welchen – etwa über eine Registrierung für die Nutzung der Services – entgeltliche Verträge abgeschlossen wurden, ist eine Gewährleistung für KI-Ergebnisse entsprechend der Anbieter von Wissens- oder Auskunftsdatenbanken denkbar. Insofern können die möglichen Folgen fehlerhafter KI-Ergebnisse einschätzbar nach vorhandenen Regelungen eingeordnet werden.

Fraglich ist jedoch, inwiefern die Haftung für KI-Ergebnisse geregelt werden kann, wenn zu den Geschädigten keine vertraglichen Beziehungen bestehen. Typisch für die Haftungsrisiken bei KI sind das Autonomierisiko und das Opazitätsrisiko. Bei ersterem können nicht vorhersehbare autonome Entscheidungen und fehlende Reproduzierbarkeit zu Schwierigkeiten bei der Ermittlung des schadensverursachenden Ereignisses führen. Das Opazitätsrisiko beschreibt die Herausforderung, die Ursache der fehlerhaften Entscheidung in einem schwer durchschaubaren Entscheidungssystem zuzuordnen. Ursache von Fehlern können sowohl in den Trainingsdaten, der Programmierung der KI und der angewandten Methodik, aber auch in gezielten Angriffen auf Trainingsdaten und/oder Anwendungen liegen. Neben dem KI-Hersteller können Datenlieferanten, KI-Anbieter (welche nicht Hersteller sind), aber auch die Nutzer Verursachungsbeiträge für fehlerhafte KI-Ergebnisse zu verantworten haben. Fehlerbeiträge der Datenlieferanten können zum einen in der Datenqualität liegen, d.h. der Aktualität und Richtigkeit der Daten, sowie der möglichen Beschreibung der Datensätze oder der Lieferung von Metadaten, soweit dies geschuldet wird.

De lege lata kommen bei fehlerhaften KI-Ergebnissen Ansprüche nach dem Produkthaftungsgesetz und nach der Produzentenhaftung gemäß § 823 Abs. 1 BGB in Betracht. Als Kriterien der Pflichtverletzung und des Sorgfaltsmaßstabs nach § 823 Abs. 1 BGB kommen die Frage des Risikos des KI-Einsatzes im Verhältnis zum menschlichen Verhalten in Betracht, neben der Beachtung des Maßstabs des Standes der anwendbaren Technik[182]. Im Rahmen der Produzentenhaftung besteht die Verpflichtung des Herstellers, für Konstruktionsfehler, Fabrikationsfehler, Instruktionsfehler und Produktbeobachtungsfehler

[178] *Hacker*, GRUR 2020, 1025 (1028); *Gräfe/Kahl*, MMR 2021, 121 (123).
[179] Zum Ausscheiden des patentrechtlichen Schutzes bei Textgeneratoren-KI: *Gräfe/Kahl*, MMR 2021, 121 (123).
[180] *Hacker*, GRUR 2020, 1025 (1031).
[181] Weitere Beispiele bei *Zech*, ZfPW 2019, 204
[182] Siehe zur Haftung für Trainingsdaten im Speziellen: *Zech*, NJW 2022, 502 (506 ff.).

einzustehen. Eine Zurechnung der Tätigkeit der KI nach § 278 BGB analog in vertraglichen Konstellationen kommt mangels menschlichen Handelns nicht in Betracht, wohingegen eine analoge Anwendung von/des § 831 BGB bei deliktischer Haftung diskutiert wird.[183] Eine Ausweitung der Haftung des § 830 Abs. 1 S. 2 BGB weg von der kumulativen zur alternativen Kausalität stellt eine vorgeschlagene Option dar, diskutiert werden daneben eine analoge Anwendung des § 7 StVG oder der Tierhalterhaftung nach § 833 S. 1 BGB.[184]

3.5 Bevorstehende Regulierungen

Die Aufnahme großer Datenmengen in KI-Systeme wird sich durch die Integration von Chatbots in Suchmaschinen und sonstige Kommunikationssysteme – auch unternehmensintern – verstärken. Ähnlich wie dies bereits bei dem Übersetzungstool DeepL zu beobachten ist, wird hierbei die Rückkopplung der KI mit Auswahlentscheidungen der Nutzer zum einen zu einem größeren Trainingsdatenpool führen, zum anderen wird sich die Bedeutung der KI für Plattformangebote wie Suchmaschinen und Onlineshops sowie Social-Media-Angebote erhöhen.
Daneben wird die Schutzfähigkeit von KI-Ergebnissen und die Einführung von Leistungsschutzrechten diskutiert[185]. Für den Umgang mit großen Datenmengen ist auch die Verbreitung und der Einsatz der KI-Ergebnisse in der Plattformwirtschaft bedeutend. Hier stehen auf EU-Ebene verschiedene Neuregulierungen für den Umgang mit großen Datenmengen durch KI an, zu nennen sind neben dem KI-VO-E die KI-Haftungsverordnung[186] (im Folgenden: „KI-HVO-E“).

3.5.1 Der Entwurf der KI-Verordnung

Der KI-VO-E sieht zum einen anwendungsorientierte Verbote bestimmter KI-Einsätze vor, zum anderen werden bestimmte KI-Anwendungen risikoabhängig reguliert. Verboten sind nach Art. 5 Abs. 1 lit. a des Entwurfs der KI-Verordnung im Bereich des „Social Scoring“[187] „das Inverkehrbringen, die Inbetriebnahme oder die Verwendung eines KI-Systems, das Techniken der unterschwelligen Beeinflussung außerhalb des Bewusstseins einer Person oder absichtlich manipulative oder täuschende Techniken einsetzt, mit dem Ziel oder der Folge, dass das Verhalten einer Person oder einer Gruppe von Personen, indem die Fähigkeit der Person, eine fundierte Entscheidung zu treffen, spürbar beeinträchtigt wird, wodurch die Person veranlasst wird, eine Entscheidung zu treffen, die sie andernfalls nicht getroffen hätte, und zwar in einer Weise, die dieser Person, einer anderen Person oder einer Gruppe von Personen erheblichen Schaden zufügt oder zufügen kann“. Ausgenommen werden sollen nach Art. 5 Abs. 1 lit. a S. 2 des Entwurfs des EU-Parlaments KI-Systeme, „die für anerkannte therapeutische Zwecke auf der Grundlage einer ausdrücklichen, nach Aufklärung erteilten Einwilligung der ihnen ausgesetzten Personen oder gegebenenfalls ihres gesetzlichen Vertreters verwendet werden sollen“. Soweit also große Datenmengen dazu benutzt werden, immer komplexere Vorhersagen für das Verhalten von Personen zu treffen, wäre die entsprechende Anwendung untersagt, wenn dies zu den oben genannten Konsequenzen führen kann. Weiterhin unzulässig sind im KI-Kontext nach Art. 5 Abs. 1 lit. ba) des Parlamentsentwurfs „das Inverkehrbringen, die Inbetriebnahme oder die Verwendung von Systemen zur biometrischen Kategorisierung, die natürliche Personen nach sensitiven oder geschützten Attributen oder Merkmalen

[183] *Burchardi*, EuZW 2022, 685 (687) m.w.N.
[184] *Zech*, NJW 2022, 502 (506 ff.).
[185] *Chiampi Ohly*, SoftwareRecht: Von der Entwicklung zum Export, S. 65.
[186] Vorschlag für eine Richtlinie des Europäischen Parlaments und des Rates zur Anpassung der Vorschriften über außervertragliche zivilrechtliche Haftung an künstliche Intelligenz (Richtlinie über KI-Haftung), COM (2022) 496 final.
[187] Dazu *Ebert/Spiecker gen. Döhmann*, NVwZ 2021, 1188 (1189).

oder auf der Grundlage von Rückschlüssen auf diese Attribute oder Merkmale kategorisieren. Dieses Verbot gilt nicht für KI-Systeme, die für anerkannte therapeutische Zwecke auf der Grundlage einer ausdrücklichen, nach Aufklärung erteilten Einwilligung der ihnen ausgesetzten Personen oder gegebenenfalls ihres gesetzlichen Vertreters verwendet werden sollen".

Einige Vorgaben des KI-VO-E betreffen in besonderem Maße den Einsatz von KI bei der Nutzung großer Datenmengen: So etwa das Verbot des Art. 5 Abs. 1 lit. c) KI-VO-E, welcher das Inverkehrbringen, die Inbetriebnahme oder die Verwendung von KI-Systemen zur Bewertung des sozialen Verhaltens oder Klassifizierung natürlicher Personen oder Gruppen von natürlichen Personen über einen bestimmten Zeitraum auf der Grundlage ihres sozialen Verhaltens oder bekannter, abgeleiteter oder vorhergesagter persönlicher Eigenschaften oder Persönlichkeitsmerkmale betrifft, wobei die soziale Bewertung zu einem oder beiden der folgenden Ergebnisse führt:
(i) Schlechterstellung oder Benachteiligung bestimmter natürlicher Personen oder ganzer Gruppen natürlicher Personen in sozialen Zusammenhängen, die in keinem Zusammenhang zu den Umständen stehen, unter denen die Daten ursprünglich erzeugt oder erfasst wurden;
(ii) Schlechterstellung oder Benachteiligung bestimmter natürlicher Personen oder ganzer Gruppen natürlicher Personen in sozialen Zusammenhängen, die in keinem Zusammenhang zu den Umständen stehen, unter denen die Daten ursprünglich erzeugt oder erfasst wurden; (iii) Schlechterstellung oder Benachteiligung bestimmter natürlicher Personen oder ganzer Gruppen natürlicher Personen, in einer Weise, die im Hinblick auf ihr soziales Verhalten oder dessen Tragweite ungerechtfertigt oder unverhältnismäßig ist.

Der KI-VO-E reguliert in besonderem Maße als hochriskant einzustufende Systeme, sodass hier die rechtskonforme Auswahl der Datensätze von besonderer Bedeutung ist. Nach Erwgr. 28 S. 4 KI-VO-E sind „das Ausmaß der negativen Auswirkungen des KI-Systems auf die durch die Charta geschützten Grundrechte (…) bei der Einstufung eines KI-Systems als hochriskant von besonderer Bedeutung". Erwgr. 28a KI-V-E des Parlaments weist darauf hin, dass „Das Ausmaß der negativen Auswirkungen des KI-Systems auf die durch die Charta geschützten Grundrechte […] bei der Einstufung eines KI-Systems als hochriskant von besonderer Bedeutung [ist]. Zu diesen Rechten gehören die Würde des Menschen, die Achtung des Privat- und Familienlebens, der Schutz personenbezogener Daten, die Freiheit der Meinungsäußerung und die Informationsfreiheit, die Versammlungs- und Vereinigungsfreiheit, die Nichtdiskriminierung, die Gleichstellung der Geschlechter, das Recht auf Bildung, der Verbraucherschutz, die Arbeitnehmerrechte, die Rechte von Menschen mit Behinderungen, Geschlechtergleichstellung, Rechte des geistigen Eigentums, das Recht auf einen wirksamen Rechtsbehelf und ein faires Gerichtsverfahren, die Unschuldsvermutung und das Verteidigungsrecht sowie das Recht auf eine gute Verwaltung". Bei der Verarbeitung großer Datenmengen kommen insbesondere die Klassifikationen im Bereich des Bildungs-, Justiz- oder Medizinsystems in Betracht (insbesondere in Auswahl- oder Bewertungsszenarien[188]). Abs. 1 Nr. 6 lit. g) des Anhangs III des KI-VO-E weist für den Bereich der Strafverfolgung auf KI-Systeme hin, die bestimmungsgemäß zur Kriminalanalyse natürlicher Personen eingesetzt werden sollen und es den Strafverfolgungsbehörden ermöglichen, große komplexe verknüpfte und unverknüpfte Datensätze aus verschiedenen Datenquellen oder in verschiedenen Datenformaten zu durchsuchen, um unbekannte Muster zu erkennen oder verdeckte Beziehungen in den Daten aufzudecken. Die Klassifikation als Hochrisiko-KI-System führt zur Notwendigkeit der Erfüllung einer Vielzahl von Anforderungen aus Kapitel 2 (Art. 8-15) und Kapitel 3 (Art. 16-29), u.a. von Transparenzanforderungen an den Einsatz der KI.[189] Unabhängig von der Einordnung als Hochrisiko-KI-System haben nach Art. 52 Abs. 1 KI-VO-E die Anbieter sicherzustellen,

[188] Beispiele bei *Molavi Vasse'i*, K&R 2022, Beil. 1 zu H. 7/8, 8.
[189] *Molavi Vasse'i*, K&R 2022, Beil. 1 zu H. 7/8, 8, 9.

dass KI-Systeme, die für die Interaktion mit natürlichen Personen bestimmt sind, so konzipiert und entwickelt werden, dass natürlichen Personen mitgeteilt wird, dass sie es mit einem KI-System zu tun haben, es sei denn, dies ist aufgrund der Umstände und des Kontexts der Nutzung offensichtlich.

Der Vorschlag des Rates vom 25.11.2022[190] definiert in seinem Art. 3 Abs. 1 lit. b als „KI-System mit allgemeinem Verwendungszweck" ein KI-System, das – unabhängig davon, wie es in Verkehr gebracht oder in Betrieb genommen wird, auch in Form quelloffener Software – vom Anbieter dazu vorgesehen ist, allgemein anwendbare Funktionen wie Bild- oder Spracherkennung, Audio- und Videogenerierung, Mustererkennung, Beantwortung von Fragen, Übersetzung und Sonstiges auszuführen; dabei kann ein KI-System mit allgemeinem Verwendungszweck in einer Vielzahl von Kontexten eingesetzt und in eine Vielzahl anderer KI-Systeme integriert werden. Diese Definition würde dann auch unspezifische KI wie ChatGPT betreffen, sodass nach Titel Ia des KI-VO-E des Rates auch solche Systeme ausdrücklich den Hochrisiko-Systemen unterfallen können (Art. 4b Abs. 1 KI-VO-E des Rates), es sei denn der Anbieter hat in den Gebrauchsanweisungen oder in den Begleitdokumenten des KI-Systems mit allgemeinem Verwendungszweck ausdrücklich jegliche Verwendung mit hohem Risiko ausgeschlossen (Art. 4c Abs. 1 KI-VO-E des Rates). Allerdings erfolgt ein solcher Ausschluss nach Art. 4c Abs. 2 KI-VO-E des Rates in gutem Glauben und gilt nicht als gerechtfertigt, wenn der Anbieter hinreichende Gründe für die Annahme hat, dass es zu einer Fehlanwendung des Systems kommen könnte. Stellt der Anbieter eine Fehlanwendung auf dem Markt fest oder wird darüber informiert, so ergreift er alle erforderlichen und verhältnismäßigen Maßnahmen, um eine weitere Fehlanwendung zu verhindern, wobei er insbesondere dem Umfang der Fehlanwendung und der Schwere der damit zusammenhängenden Risiken Rechnung trägt (Art. 4c Abs. 3 KI-VO-E des Rates).

Das EU-Parlament weist in Erwgr. 60e) auf die Bedeutung großer Datenmengen für solche KI mit allgemeinem Verwendungszweck hin: „Basismodelle sind eine neuere Entwicklung, bei der KI-Modelle auf der Grundlage von Algorithmen entwickelt werden, die im Hinblick auf Allgemeinheit und Vielseitigkeit der Ergebnisse optimiert wurden. Diese Modelle werden häufig auf der Grundlage eines breiten Spektrums von Datenquellen und großer Datenmengen trainiert, um eine Fülle nachgelagerter Aufgaben zu erfüllen, darunter auch solche, für die sie nicht speziell entwickelt und trainiert wurden. Das Basismodell kann unimodal oder multimodal sein und durch verschiedene Methoden wie überwachtes Lernen oder bestärkendes Lernen trainiert werden. KI-Systeme mit spezifischer Zweckbestimmung oder KI-Systeme mit allgemeinem Verwendungszweck können eine Implementierung eines Basismodells sein, was bedeutet, dass jedes Basismodell in unzähligen nachgelagerten KI-Systemen oder KI-Systemen mit allgemeinem Verwendungszweck wiederverwendet werden kann. Diese Modelle sind für viele nachgelagerte Anwendungen und Systeme von wachsender Bedeutung." Die Einstufung als Hochrisiko-System hat im Übrigen auch Konsequenzen für vertragliche Gewährleistungsregelungen und im deliktischen Bereich für die Bestimmung von Verkehrssicherungspflichten nach § 823 Abs. 1 BGB oder die Bestimmung von Schutzgesetzen nach § 823 Abs. 2 BGB.[191]

3.5.2 Haftungsregelungen

Der ergänzende KI-HVO-E regelt in seinem Art. 1 Abs. 1 lit. a und Art. 3 die Offenlegung von Beweismitteln betreffend Hochrisiko-KI-Systeme[192] mit dem Ziel, es einem Kläger zu

[190] *Rat der Europäischen Union,* 2021/0106(COD), https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/de/pdf.
[191] *Roos/Weitz*, MMR 2021, 844 (850).
[192] Siehe zur KI-Terminologie der KI-Haftungsrichtlinie und zum Ratsvorschlag vom 3. 11. 2022 (Nr. 13955/22) *Bomhard/Siglmüller*, RDi 2022, 506 (507).

ermöglichen, einen außervertraglichen[193] verschuldensabhängigen zivilrechtlichen Schadensersatzanspruch zu begründen sowie die Beweislast bei der Geltendmachung außervertraglicher verschuldensabhängiger zivilrechtlicher Ansprüche vor nationalen Gerichten in Bezug auf Schäden, die durch ein KI-System verursacht wurden. Nach Art. 4 KI-HVO-E wird eine widerlegbare Vermutung eines ursächlichen Zusammenhangs im Fall eines Verschuldens eingeführt. Diese greift nach Art. 4 Abs. 2 KI-HVO-E, wenn bei dem KI-System Techniken eingesetzt werden, bei denen Modelle mit Daten trainiert werden, und das System nicht anhand von Trainings-, Validierungs- und Testdatensätzen entwickelt wurde, die den in Art. 10 Abs. 2-4 des Kommissionsentwurfs zur KI-VO genannten Qualitätskriterien entsprechen. Neben dem KI-HVO-E ist bei der künftigen Regulierung der Entwurf der ProduktHaftRL[194] (im Folgenden: „PHL-E") zu berücksichtigen:

Nach Erwgr. 3 S. 1 PHL-E soll die bisherige Richtlinie 85/374/EWG „vor dem Hintergrund der Entwicklungen im Zusammenhang mit neuen Technologien, einschließlich künstlicher Intelligenz (KI), neuer Geschäftsmodelle der Kreislaufwirtschaft und neuer globaler Lieferketten, die zu Inkonsistenzen und Rechtsunsicherheit insbesondere in Bezug auf die Bedeutung des Begriffs „Produkt" geführt haben, überarbeitet werden". Nach S. 4 des Erwgr. 3 wird die Überarbeitung „daher die Bereitstellung und Nutzung solcher neuen Technologien, einschließlich KI, fördern und gleichzeitig sicherstellen, dass Kläger unabhängig von der betreffenden Technologie von demselben Schutzniveau profitieren können".
Nach Art. 4 Abs. 3 bis 5 PHL-E werden neben dem Produkt auch folgende Elemente definiert, welche im KI-Kontext von Relevanz sind:
- „Komponente" bezeichnet jeden materiellen oder immateriellen Gegenstand und jeden verbundenen Dienst, der vom Hersteller eines Produkts oder unter Kontrolle des Herstellers in das Produkt integriert oder mit dem Produkt verbunden wird.
- „Verbundener Dienst" bezeichnet einen digitalen Dienst, der so in ein Produkt verbunden ist, dass das Produkt ohne ihn eine oder mehrere seiner Funktionen nicht ausführen könnte;
- „Kontrolle des Herstellers" bezeichnet die Tatsache, dass der Hersteller eines Produkts a) die Integration, Verbindung oder Lieferung einer Komponente einschließlich Software-Updates oder -Upgrades durch einen Dritten oder b) die Änderung des Produkts genehmigt.

Als Hersteller des Produkts gilt nach Art. 7 Nr. 4 PHL-E dabei auch jede natürliche oder juristische Person, die ein bereits in Verkehr gebrachtes oder in Betrieb genommenes Produkt verändert, wenn die Änderung nach den einschlägigen Vorschriften des Unions- oder des nationalen Rechts über die Produktsicherheit als wesentlich gilt und außerhalb der Kontrolle des ursprünglichen Herstellers erfolgt. Dementsprechend können die Herstellung und Verbreitung eines Produkts, welches KI-Anteile enthält, dazu führen, dass auch der Hersteller der integrierten KI bei entsprechenden wesentlichen Änderungen der Funktionalität des Produkts durch die KI als Hersteller im Sinne des PHL-E gilt. Hinsichtlich der Beweislast ist nach Art. 9 Abs. 1 PHL-E durch die Mitgliedstaaten sicherzustellen, dass der Kläger verpflichtet ist, die Fehlerhaftigkeit des Produkts, den erlittenen Schaden und den ursächlichen Zusammenhang zwischen der Fehlerhaftigkeit und dem Schaden nachzuweisen. Allerdings wird nach Art. 9 Abs. 2 PHL-E von der Fehlerhaftigkeit des Produkts unter anderem dann ausgegangen, wenn der Kläger nachweist, dass der Schaden durch eine offensichtliche Funktionsstörung des Produkts bei normaler Verwendung oder unter normalen Umständen verursacht wurde. Nach Art. 9 Abs. 3 wird von einem ursächlichen

[193] *Bomhard/Siglmüller*, RDi 2022, 506 (507).
[194] Vorschlag für eine Richtlinie des europäischen Parlaments und des Rates über die Haftung für fehlerhafte Produkte, COM/2022/495 final.

Zusammenhang zwischen der Fehlerhaftigkeit des Produkts und dem Schaden ausgegangen, wenn festgestellt wurde, dass das Produkt fehlerhaft und der entstandene Schaden von der dem betreffenden Fehler typischerweise entsprechenden Art ist.

3.5.3 Digital Markets Act, Digital Services Act und Data Act

Soweit Ergebnisse einer KI, die große Datenmengen verarbeitet, in sehr große Suchmaschinen und sehr große zentrale Plattformdienste eingebunden werden,[195] können die Nutzungen – wenn sie die weiteren Anwendungsvoraussetzungen der jeweiligen Rechtsakte erfüllen – unter die Regelungen des Digital Services Act[196] (DSA) und/oder des Digital Market Act (DMA) fallen. Der Einsatz solcher KI-Systeme muss nach Art. 34 Abs. 1 DSA einer Risikobewertung unterzogen werden, wenn es sich um Anbieter sehr großer Online-Plattformen oder sehr großer Online-Suchmaschinen handelt.[197] Diese müssen sorgfältig alle systemischen Risiken in der Union ermitteln, so unter anderem nach Art. 34 Abs. 1 S. 3 DSA systemische Risiken wie die Verbreitung rechtswidriger Inhalte, etwaige tatsächliche oder vorhersehbare nachteilige Auswirkungen auf die Ausübung der Grundrechte, tatsächliche oder absehbare nachteilige Auswirkungen auf die gesellschaftliche Debatte und auf Wahlprozesse und die öffentliche Sicherheit und weitere grundrechtsrelevante Risiken. Der Digital Markets Act[198] regelt unter anderem die Verpflichtung von digitalen Torwächtern, welche ein Ranking auf Plattformen anhand transparenter, fairer und diskriminierungsfreier Bedingungen vornehmen (Art. 6 Abs. 5 S. 2 DMA). Soweit KI, welche große Datenmengen verarbeitet, in das Ranking maßgeblich eingebunden wird, sind die entsprechenden Vorgaben des DMA daher einschlägig und ebenfalls zu beachten[199]. Zu beachten sind weiterhin die Regelungen des sich im Entwurfsstadium[200] befindlichen Data Acts (im Folgenden: „DA-E"). Der Data Act soll die Möglichkeit des Zugangs und der sektorenübergreifenden Verwertung von Daten auf vernetzten Geräten schaffen. Im Hinblick auf den Einsatz von KI und die Verarbeitung großer realer Datenmengen sind einige einschlägige Vorschriften zu beachten, welche sich auf Konzeption und Betrieb der entsprechenden Systeme auswirken.

So regelt Art. 1 Abs. 1 lit. a) des DA-E unter anderem die Gestaltung vernetzter Produkte, um dem Nutzer eines vernetzten Produkts Zugang zu den Daten zu ermöglichen, die von diesem vernetzten Produkt oder während der Erbringung verbundener Dienste erzeugt werden. Nach Art. 2 Nr. 4 DA-E wird als „virtuelle Assistenten" Software definiert, die Aufträge, Aufgaben oder Fragen verarbeiten kann, auch aufgrund von Eingaben in Ton- und Schriftform, Gesten oder Bewegungen, und auf der Grundlage dieser Aufträge, Aufgaben oder Fragen den Zugang zu anderen Diensten gewährt oder die Funktionen von Produkten steuert. Art. 2 Nr. 6 DA-E beschreibt als „Dateninhaber" eine juristische oder natürliche Person, die auf Daten aus dem vernetzten Produkt zugegriffen oder bei der Erbringung eines verbundenen Dienstes Daten erzeugt hat und die das vertraglich vereinbarte Recht hat, diese Daten zu nutzen, und die gemäß dieser Verordnung, dem geltenden Unionsrecht oder den nationalen Rechtsvorschriften zur Umsetzung des Unionsrechts

[195] Artikel von *Weiß* auf heise.de vom 14. 2. 2023, „Klar macht das neue Bing Fehler – spannend ist die KI-Suche dennoch", https://www.heise.de/news/Klar-macht-das-neue-Bing-Fehler-spannend-ist-die-KI-Suche-dennoch-7493727.html.

[196] VO (EU) 2022/2065 des Europäischen Parlaments und des Rates vom 19. 10. 2022 über einen Binnenmarkt für digitale Dienste und zur Änderung der RL 2000/31/EG (Gesetz über digitale Dienste).

[197] Diese müssen nach Art. 33 Abs. 1 DSA eine durchschnittliche monatliche Zahl von mindestens 45 Millionen aktiven Nutzern in der Union haben und gemäß Art. 33 Abs. 4 DSA durch die Kommission als solche benannt werden.

[198] VO (EU) 2022/1925 des Europäischen Parlaments und des Rates vom 14. 9. 2022 über bestreitbare und faire Märkte im digitalen Sektor und zur Änderung der Richtlinien (EU) 2019/1937 und (EU) 2020/1828 (Gesetz über digitale Märkte), ABl. 2022, L 265/1.

[199] Zum Verhältnis zum KI-VO-E *Hacker*, GRUR 2022, 1278 (1279).

[200] Rat und Parlament haben am 27.06.2023 dazu eine vorläufige Einigung erzielt, siehe https://www.consilium.europa.eu/de/press/press-releases/2023/06/27/data-act-council-and-parliament-strike-a-deal-on-fair-access-to-and-use-of-data/.

verpflichtet ist, dem Nutzer oder einem Datenempfänger bestimmte Daten zur Verfügung zu stellen.

Art. 3 Abs. 1 DA-E legt als Pflicht fest, dass vernetzte Produkte so konzipiert und hergestellt werden, dass auf von ihnen erhobene, erzeugte oder anderweitig erhaltene Daten für den Nutzer standardmäßig kostenlos und einfach sicher und – soweit relevant und technisch machbar – in einem umfassenden, strukturierten, gängigen und maschinenlesbaren Format unmittelbar zugänglich sind. Die Daten müssen in der Form zur Verfügung stehen, in der sie von dem vernetzten Produkt erhoben, erhalten oder erzeugt wurden, wobei nur die minimalen Anpassungen vorgenommen werden, die erforderlich sind, um sie für Dritte nutzbar zu machen, einschließlich der zugehörigen Metadaten, die zur Interpretation und Nutzung der Daten benötigt werden. Diese Vorgabe dürfte bei Daten, welche Informationen, die mithilfe komplexer proprietärer Algorithmen aus diesen Daten abgeleitet bzw. gefolgert wurden, schwer umzusetzen sein[201]. Allerdings sieht Art. 4 Abs. 1 S. 3 DA-E vor, dass Informationen, die mithilfe komplexer proprietärer Algorithmen aus diesen Daten abgeleitet bzw. gefolgert wurden, nicht unter die Verpflichtung des Dateninhabers zur Weitergabe von Daten an Nutzer bzw. Datenempfänger fallen, sofern der Nutzer und der Dateninhaber nichts anderes vereinbart haben. Dies gilt insbesondere, wenn dabei die Ausgabe mehrerer Sensoren in dem vernetzten Produkt kombiniert wurde. Die Erwägungsgründe lassen offen, inwiefern damit KI-Ergebnisse eindeutig aus der Zugänglichmachungspflicht ausgenommen sind. Nach Art 5 Abs. 1 DA-E sind diese Daten auf Verlangen auch bestimmten Dritten zur Verfügung zu stellen.

3.6 Fazit

Bei der Verarbeitung großer Datenmengen durch KI stellt sich eine ganze Reihe schutzrechtlicher Herausforderungen. Zum einen ist zu klären, ob ohne vertraglich abgesicherte Datengewinnungen eine Datenerhebung aus frei zugänglichen Internetseiten de lege lata rechtlich zulässig ist. Dies hängt – neben der hier ausgeklammerten datenschutzrechtlichen Zulässigkeit – von der möglichen Schutzfähigkeit der auszulesenden Daten sowie der Anwendbarkeit der §§ 44b, 60d UrhG ab, daneben von der möglichen rechtlichen und technischen Absicherung der Datensätze durch die Internetseitenbetreiber. Im Zuge der Verarbeitung ist in der Folge sicherzustellen, dass mögliche urheberrechtlich geschützte Datensätze nicht so in die KI-Ergebnisse einfließen, dass diese eine Bearbeitung nach § 23 Abs. 1 UrhG darstellen. Haftungsrechtlich ist die Verarbeitung großer Datensätze mit der Herausforderung behaftet, dass die KI-Betreiber gegenüber Dritten de lege lata auch für mögliche Fehler in den Datensätzen haften. Im Hinblick auf den KI-VO-E ist unabhängig von der Größe der zu verarbeitenden Datensätze der Einsatzzweck für die notwendigen Maßnahmen zur Risikobeherrschung maßgeblich. Die neuen Haftungsrichtlinien der KI-HVO-E enthalten Offenlegungspflichten, welche im Hinblick auf die Opazität und Autonomie der KI gerade bei großen Datenmengen schwer umzusetzen sein werden und die Frage aufwerfen, wie die zu benennenden beweisrelevanten Daten ermittelt werden können. Die Nichterfüllung dieser Pflichten kann zu Beweislastumkehrungen führen, welche die Haftungssituation beeinflussen. Es wird daher eine Herausforderung für KI-Betreiber sein, bei der Verarbeitung großer Datenmengen sichere vertragliche Regelungen zu finden, welche sich im Einklang mit den künftigen Regulierungen befinden, ohne den eigenen Spielraum zur Verwertung der KI-Ergebnisse zu sehr einzuengen.

[201] Siehe im Einzelnen zu den Pflichten *Hennemann/Steinrötter*, NJW 2022, 1481.

4. Verifizierung der datenschutzkonformen Anonymisierung und Re-Identifizierung von Daten in großen realen Datenverarbeitungssystemen

Bei der Verarbeitung großer Datensätze kann die Eingriffsintensität in die Rechte und Freiheiten der betroffenen Personen erheblich reduziert werden, wenn personenbezogene Daten anonymisiert werden. Zudem entfällt darüber regelmäßig die Anwendbarkeit der DSGVO-Vorschriften, sodass Verarbeiter geringeren Anforderungen ausgesetzt sind.

Aufgrund der vagen Begriffsbestimmungen in der DSGVO ist es für Organisationen sehr schwer zu bestimmen, wann Daten aus rechtlicher Sicht tatsächlich anonym(isiert) sind, sodass ein erhebliches Maß an Rechtsunsicherheit besteht, wenn Organisationen anonyme/anonymisierte Daten verarbeiten wollen.[202] Das folgende Unterkapitel beschäftigt sich vor diesem Hintergrund mit der Frage der Definition und Verifikation von Anonymität. Ziel ist es aufzuzeigen, welche rechtlichen Fragen rund um die Anonymisierung personenbezogener Daten zu lösen sind, um Organisationen Rechtssicherheit über die durch sie verarbeiteten (anonymen) Daten zu geben.

4.1 Unterscheidung zwischen personenbezogenen Daten und anonymen Daten

Gemäß Art. 4 Nr. 1 DSGVO sind personenbezogene Daten alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen, wobei eine identifizierbare natürliche Person eine Person ist, die direkt oder indirekt identifiziert werden kann. Damit Daten als personenbezogene Daten gelten, ist es also nicht erforderlich, dass die betroffene Person durch direkte Kennungen wie den Namen und damit verbundene Informationen identifiziert werden kann.[203] Vielmehr genügt auch die Möglichkeit, eine natürliche Person indirekt zu identifizieren, insbesondere durch sogenannte Quasi-Identifikatoren. Quasi-Identifikatoren sind Datensätze, die aus Attributen bestehen, die – jedes Attribut für sich genommen – keine natürliche Person identifizieren, aber in ihrer Kombination identifizierend werden. Beispielsweise können Informationen über das Alter, das Geschlecht und den Wohnort in ihrer Kombination als Quasi-Identifikator und je nach Kontext als indirekt identifizierend angesehen werden.[204]

Eine besondere Form der personenbezogenen Daten sind pseudonymisierte Daten. Der Begriff der pseudonymisierten Daten ist in der DSGVO nicht direkt definiert. Aus der Definition der Pseudonymisierung in Art. 4 Nr. 5 DSGVO lässt sich jedoch ableiten, dass pseudonymisierte Daten im Sinne der DSGVO personenbezogene Daten sind, die ohne Hinzuziehung zusätzlicher Informationen nicht mehr einer spezifischen betroffenen Person zugeordnet werden können, sofern diese zusätzlichen Informationen gesondert aufbewahrt werden und Gegenstand technischer und organisatorischer Maßnahmen sind.

---

[202] Das vorliegende Kapitel beruht in weiten Teilen auf *Stummer* 2022, S. 179 ff.
[203] *Art. 29-Datenschutzgruppe, Stellungnahme 4/2007 zum Begriff „personenbezogene Daten“*, S. 15; *Schild* in *Wolff/Brink*, BeckOK Datenschutzrecht, Art. 4 DSGVO Rdnr. 16; *Karg* in *Simitis/Hornung/Spiecker gen. Döhmann*: Datenschutzrecht, Art. 4 Rdnr. 1 ff.
[204] *Sweeney, Simple Demographics Often Identify People Uniquely, S.7.*

Den Begriff der anonymen Daten lässt der europäische Verordnungsgeber undefiniert. Stattdessen wird der Begriff nur in den Erwägungsgründen der DSGVO erwähnt, insbesondere in Erwgr. 26 DSGVO. Demnach gelten die Grundsätze des Datenschutzes nur für Informationen über eine identifizierte oder identifizierbare natürliche Person und sind daher nicht auf anonyme Daten anwendbar. Anonyme Daten unterscheiden sich von personenbezogenen Daten insofern, als anonyme Informationen sich nicht auf eine identifizierte oder identifizierbare natürliche Person beziehen. Anonyme Daten können aber auch (ehemals) personenbezogene Daten sein, die anonymisiert wurden, sodass die betroffene Person nicht mehr identifizierbar ist. Anonyme Daten können daher als das Gegenteil von personenbezogenen Daten verstanden werden.[205]

Ausgehend von den rechtlichen Definitionen der Begriffe „anonyme Daten" und „personenbezogene Daten" wird deutlich, dass sich die Begriffe in ihrer Fähigkeit, eine natürliche Person zu identifizieren, sowie in ihrem Grad der Identifizierbarkeit unterscheiden. In dieser Hinsicht scheinen die Grenzen zwischen den Begriffen klar zu sein: Personenbezogene Daten ermöglichen im Allgemeinen die Identifizierung einer natürlichen Person und weisen daher einen höheren Grad der Identifizierbarkeit auf als anonyme Daten, bei denen die Identifizierung einer bestimmten natürlichen Person nicht (mehr) möglich ist. Bei den personenbezogenen Daten muss jedoch zusätzlich unterschieden werden zwischen „direkt identifizierenden personenbezogenen Daten" und „indirekt identifizierenden personenbezogenen Daten". Da direkt identifizierende personenbezogene Daten die unmittelbare Identifizierung einer natürlichen Person ermöglichen, während indirekt identifizierende personenbezogene Daten die Identifizierung nur in Verbindung mit anderen Informationen ermöglichen und anonyme Daten überhaupt keine Identifizierung zulassen, steigt der Grad der Identifizierbarkeit von anonymen Daten über indirekt identifizierende personenbezogene Daten zu direkt identifizierenden personenbezogenen Daten.

4.2 Beurteilung, ob es sich bei Daten um personenbezogene Daten oder anonyme Daten handelt

Obwohl die Definitionen und die Unterschiede zwischen den Begriffen klar zu sein scheinen, verschwimmen die Grenzen zwischen ihnen, wenn es darum geht zu beurteilen, ob sich Daten auf eine identifizierte oder identifizierbare natürliche Person beziehen, ob ein Datensatz pseudonymisiert oder anonymisiert verwendet wird, und ob ein Datensatz anonym ist oder nicht. Ob personenbezogene Daten direkt oder indirekt identifizierend sind, hängt zunächst vom jeweiligen Kontext ab, insbesondere von der Größe der Gruppe der in Frage kommenden natürlichen Personen („Anonymitätsmenge" oder „Anonymitätsgruppengröße"[206]).[207] So sind Namen bezogen auf die Weltbevölkerung regelmäßig nicht einzigartig und ermöglichen somit i.d.R. keine direkte Identifizierung. Innerhalb einer Fußballmannschaft, eines Unternehmens oder einer Schulklasse können Namen hingegen eine direkte Identifizierung ermöglichen.[208] Daher muss bei der Beurteilung, ob Daten direkt oder indirekt identifizierend sind, der Kontext der Verarbeitung, insbesondere im Hinblick auf die Anonymität, berücksichtigt werden.[209] Eine Einstufung von Daten als direkt identifizierende personenbezogene Daten oder indirekt identifizierende personenbezogene Daten allein auf der Grundlage der Daten ist nicht ausreichend.

[205] *Roßnagel*, ZD 2021, 188 (189); *Karg* in *Simitis/Hornung/Spiecker gen. Döhmann*: Datenschutzrecht, Art. 4 Rdnr. 19 ff.; *Schreiber/Stummer* in *Selzer*: Datenschutzrecht, Art. 4 Rdnr. 7.
[206] *Pfitzmann/Köhntopp/Shostack, Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology*, https://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.22.pdf, S. 4.
[207] *Art. 29-Datenschutzgruppe, Stellungnahme 4/2007 zum Begriff „personenbezogene Daten"*, S. 12 ff.
[208] Vgl. *Art. 29-Datenschutzgruppe, Stellungnahme 4/2007 zum Begriff „personenbezogene Daten"*, S. 15.
[209] *Article 29 Data Protection Working Party*, *Opinion 4/2007 on the concept of personal data*, S.12ff.; *Karg* in *Simitis/Hornung/Spiecker gen. Döhmann*: Datenschutzrecht, Art. 4 Rdnr. 22f.

Noch komplexer wird es, wenn es um die Frage geht, ob es sich bei den Daten überhaupt um personenbezogene Daten handelt oder ob sie als anonyme Daten einzustufen sind. Insbesondere die Grenze zwischen anonymen Daten und *identifizierbaren* personenbezogenen Daten ist fließend. Dies liegt daran, dass gemäß Erwgr. 26 DSGVO bei der Beurteilung, ob eine natürliche Person identifizierbar ist, alle Mittel berücksichtigt werden, die nach vernünftigem Ermessen von dem für die Verarbeitung Verantwortlichen oder einer anderen Person zur Identifizierung der natürlichen Person genutzt werden können. Die Wahrscheinlichkeit der eingesetzten Mittel hängt von allen objektiven Faktoren ab, wie den Kosten der Identifizierung, dem dafür erforderlichen Zeitaufwand und den verfügbaren Identifizierungstechnologien. Anstelle einer klaren Unterscheidung zwischen (indirekt identifizierenden) personenbezogenen Daten und anonymen Daten sieht der europäische Verordnungsgeber somit verschiedene Faktoren vor, die bei der Entscheidung, ob Daten einen Bezug zu einer natürlichen Person aufweisen oder anonym sind, berücksichtigt werden müssen. Infolgedessen ist es aus rechtlicher Sicht schwer eindeutig zu bewerten, wann Daten anonym sind, da der Grat zwischen (indirekt identifizierenden) personenbezogenen Daten und anonymen Daten ein sehr schmaler ist und unklar ist, wo genau die Grenze zwischen den beiden Begriffen verläuft.

Um Rechtssicherheit für datenanonymisierende Organisationen zu schaffen und sie in die Lage zu versetzen, die Anonymisierung und ihr Wertschöpfungspotenzial zu nutzen, muss daher konkretisiert werden, wo die Grenze zwischen personenbezogenen (bzw. personenbeziehbaren) Daten und anonymen Daten verläuft, welche Voraussetzungen erfüllt sein müssen, um Daten als hinreichend anonym zu betrachten, und unter welchen Bedingungen Daten keine personenbezogenen Daten sind, sondern als anonym zu qualifizieren sind.

Dementsprechend gilt es zur Erreichung von Rechtssicherheit bei der Datenanonymisierung folgende Fragen zu lösen:

- Wo liegt die Grenze zwischen anonymen Daten und personenbezogenen Daten? Wann sind Daten anonym?
- Welche Voraussetzungen müssen erfüllt sein, damit Daten als ausreichend anonym gelten?

4.3 Relevanz der Relativität von Anonymität

Fraglich ist zudem, ob Anonymität absolut oder relativ sein muss.

Nach der absoluten oder objektiven Theorie der Identifizierbarkeit können Daten nur dann als anonym gelten, wenn es für niemanden möglich ist, die betroffene natürliche Person zu identifizieren. Die relative oder subjektive Theorie der Identifizierbarkeit hingegen betrachtet Daten als anonym, wenn eine Identifizierung aufgrund aller Faktoren in der konkreten Situation, wie z. B. der Kosten und des Zeitaufwands für die Identifizierung, vernünftigerweise nicht wahrscheinlich ist (auch wenn sie theoretisch möglich wäre).[210]

In Bezug auf diesen Streit ist die DSGVO nicht unmissverständlich. Erwgr. 26 DSGVO enthält sowohl Elemente der absoluten als auch der relativen Theorie: Für die relative Theorie der Identifizierbarkeit spricht, dass bei der Feststellung, ob eine natürliche Person identifizierbar ist, nur die Mittel berücksichtigt werden sollten, die nach vernünftigem Ermessen für die Identifizierung der natürlichen Person verwendet werden können. Andererseits bezieht sich die DSGVO auf alle Mittel, die von dem für die Verarbeitung Verantwortlichen

[210] *Karg* in *Simitis/Hornung/Spiecker gen. Döhmann*: Datenschutzrecht, Art. 4 Rdnr. 48 ff.; *Schreiber/Stummer* in *Selzer*: Datenschutzrecht, Art. 4 Rdnr. 6.

oder einer anderen Person verwendet werden, was für die absolute Theorie der Identifizierbarkeit charakteristisch ist.
Der Europäische Gerichtshof verfolgt zur Beurteilung des Personenbezugs und der Anonymität im Grundsatz einen relativen Ansatz: Laut Europäischen Gerichtshofs sind bei der Beurteilung, ob Daten personenbezogen oder anonym sind, alle rechtlichen Mittel zu berücksichtigen, die es einer Partei ermöglichen, eine natürliche Person zu identifizieren[211] (einschließlich der Einbeziehung von Dritten, die über die Informationen zur Identität verfügen und rechtlich gezwungen sind, diese Informationen bereitzustellen).[212] Daraus folgt, dass nur die rechtmäßigen Mittel der Parteien, die Zugriff auf den anonymisierten Datensatz haben, berücksichtigt werden müssen.

Die Frage, ob und inwieweit Anonymität absolut oder relativ sein muss, ist noch nicht abschließend beantwortet. Dementsprechend gilt es zur Erreichung von Rechtssicherheit bei der Datenanonymisierung folgende Fragen zu lösen:

- Ist bei der Beurteilung der Anonymität ein relativer oder absoluter Ansatz anzuwenden?
- (Sofern ein relativer Ansatz anzuwenden ist): Wo verläuft die Grenze der relativen Beurteilung von Anonymität?

4.4 Auswirkungen einer möglichen De-Anonymisierung

In der Vergangenheit gab es mehrere Beispiele, bei denen die vermeintliche Anonymität umgekehrt wurde (sog. „De-Anonymisierung" oder „Re-Identifikation"), sodass Rückschlüsse auf eine bestimmte natürliche Person (wieder) möglich waren. Ein Beispiel für die Re-Identifizierbarkeit von Daten liefern u.a. *de Montjoye et al*,[213] die Geolokationsdaten von 1,5 Millionen Personen über einen Zeitraum von 15 Monaten beobachteten und herausfanden, dass es nur vier Standort-Zeit-Punkte braucht, um 95 % der beobachteten Personen eindeutig zu identifizieren.

Eine solche De-Anonymisierung kann durch naiv oder unzureichend angewandte Anonymisierungstechniken,[214] durch neue, schnellere oder billigere Technologien[215] sowie durch die zunehmende Zugänglichkeit von Daten[216] verursacht werden. So wird die Möglichkeit, eine natürliche Person zu identifizieren, durch den technologischen Fortschritt und die Vernetzung über das Internet verstärkt; beides führt zu einer zunehmenden Menge und Zugänglichkeit von Daten. Dies kann dazu führen, dass zunächst anonyme Daten „schleichend" personenbezogene Daten werden.[217] So werden beispielsweise zur Kontaktaufnahme mit Freunden, Familienmitgliedern oder anderen Kontakten in vielen Fällen Direktnachrichtendienste wie WhatsApp oder Instagram Direct Messaging verwendet. Darüber hinaus werden persönliche Treffen und Freizeitaktivitäten häufig von Telefonkameras aufgezeichnet und, möglicherweise in Verbindung mit einer Geolokalisierung, mit anderen Personen über die sozialen Medien geteilt. Ein weiteres Beispiel findet sich im

[211] EuGH, C-582/14, 19.10.2016.
[212] BGH, VI ZR 135/13, 16.05.2017.
[213] *De Montjoye/Hidalgo/Verleysen/Blondel*, SCIENTIFIC REPORT 2013, 1 (1).
[214] Aus technischer Sicht gibt es zahlreiche Techniken zur Anonymisierung von Daten. Grundlage für die Anonymisierung ist in der Regel die Entfernung oder Maskierung von direkten Identifikatoren, so dass eine direkte Identifizierung nicht mehr möglich ist. Eine vollständige Anonymisierung von Daten setzt aber zusätzlich voraus, dass eine indirekte Identifizierung einer natürlichen Person praktisch ausgeschlossen ist. Daher müssen zusätzliche Techniken zur Generalisierung und/oder Randomisierung von Daten so weit angewendet werden, dass natürliche Personen nicht mehr identifizierbar sind. Durch die Generalisierung werden die Daten so abstrahiert und verallgemeinert, dass die Identifizierung einer natürlichen Person unwahrscheinlicher wird. Durch Randomisierung werden die Daten so verzerrt, dass der Bezug zu einer natürlichen Person aufgehoben wird. Siehe hierzu auch: *Article 29 Data Protection Working Party*, *Opinion 05/2014 on Anonymisation Techniques*, S.12ff.
[215] *Hornung/Wagner*, CR 2019, 565 (565 ff.); *Hornung/Wagner*, ZD 2020, 223 (223 ff.).
[216] *Bohannon*, SCIENCE 2013, 262 (262).
[217] Zu diesem Effekt und seiner rechtlichen Bewertung *Hornung/Wagner*, CR 2019, 565.

Haushalt: Immer mehr Haushalte nutzen Smart-Home-Geräte, wie Staubsaugerroboter, smarte Kühlschränke, smarte Waschmaschinen und Sprachassistenten (z.B. Alexa), die unterstützen, indem sie das Licht einschalten oder Timer einstellen, aber auch nach Rezepten suchen, Fragen beantworten oder Waren bestellen. All diese und zahlreiche weitere Aktivitäten verarbeiten Daten und führen daher zu einer zunehmenden Menge und Zugänglichkeit von Daten, die Rückschlüsse auf den Einzelnen zulassen. Es stellt sich daher die Frage, ob in einer Ära von Big Data, in der durch die Nutzung von Diensten eine große Menge an Daten gesammelt wird, Informationen im Internet ausgetauscht werden und der Einzelne eine digitale Spur hinterlässt, Anonymität überhaupt noch möglich ist.

Aufgrund der drohenden Möglichkeit einer De-Anonymisierung infolge von u.a. schnelleren oder günstigeren Technologien sowie einer erhöhten Verfügbarkeit und Zugänglichkeit von Daten, die zur De-Anonymisierung genutzt werden können, sowie aufgrund der Bezugnahme des Erwägungsgrunds 26 DSGVO auf den technologischen Fortschritt besteht nach der herrschenden Meinung eine Pflicht zur kontinuierlichen Überwachung des Zustands der Anonymität.[218] Allgemein anerkannte Methoden zur Überwachung des Zustands der Anonymität existieren dabei jedoch noch nicht. Fraglich ist daher auch, wie der Zustand der Anonymität langfristig überwacht, verifiziert und sichergestellt werden kann. Ein schon länger vorgeschlagenes Instrument, nämlich rechtliche Verbote der De-Anonymisierung, gibt es bisher nur im Ausland.[219]

Dementsprechend gilt es zur Erreichung von Rechtssicherheit bei der Datenanonymisierung folgende Fragen zu lösen:

- Ist Anonymität aus rechtlicher Sicht überhaupt möglich?
- Wie kann der Zustand der Anonymität überwacht, überprüft und langfristig gesichert werden?

4.5 (Rechtliche) Konsequenzen der De-Anonymisierung

Darüber hinaus sind die Konsequenzen der De-Anonymisierung zu prüfen. Auch wenn klar ist, dass Daten infolge der De-Anonymisierung wieder personenbezogen werden, sodass bei deren Verarbeitung grundsätzlich die datenschutzrechtlichen Pflichten eingehalten werden müssen, ist bisher weder gerichtlich noch durch die Aufsichtsbehörden geklärt, mit welchen konkreten rechtlichen Konsequenzen zu rechnen ist, wenn anonym geglaubte Daten de-anonymisiert werden, und wie mit dem Problem des schwer feststellbaren Zeitpunkts des Übergangs von anonymen zu personenbezogenen Daten umzugehen ist.[220]

Des Weiteren bedarf die Einhaltung mancher datenschutzrechtlicher – und im Falle der De-Anonymisierung von Daten ggf. bestehender – Pflichten (z.B. die Einholung von Einwilligungen) Kenntnis über die Kontaktmöglichkeiten der betroffenen Personen, welche ggf. auch nach einer De-Anonymisierung nicht vorliegt. Es stellt sich daher die Frage, welche Folgen eine unzureichende Anonymisierung bzw. De-Anonymisierung hat und ob und inwieweit die rechtlichen Verpflichtungen der DSGVO erfüllt werden müssen, wenn natürliche Personen wieder identifizierbar werden.

---

[218] Art. 29-Datenschutzgruppe, WP 216, S. 29; Marnau, DuD 2016, 428, 429; *Arning/Rothkegel* in *Taeger/Gabel*, Art. 4 DSGVO Rdnr. 48; *Laue/Kremer/Laue*, § 1 Rdnr. 21.

[219] Zu diesem Regulierungsansatz schon *Roßnagel/Scholz*, MMR 2000, 721 (730 f.); zu entsprechenden Vorschriften in Japan *Roßnagel/Geminn*, ZD 2021, 487 (488 ff.); zu präventiven Maßnahmen de lege lata und de lege ferenda auch *Hornung/Wagner*, CR 2019, 565 (572 ff.).

[220] Zum Problem der ausgelösten Rechtspflichten s. die Untersuchung bei *Hornung/Wagner*, CR 2019, 565 (568 ff.), zum Problem des Zeitpunkts ebd., 670 ff.; *Karg* in *Simitis/Hornung/Spiecker gen. Döhmann*: Datenschutzrecht, Art. 4 Rdnr. 19ff.; *Schreiber/Stummer* in *Selzer*: Datenschutzrecht, Art. 4 Rdnr. 2 ff.

Dementsprechend gilt es zur Erreichung von Rechtssicherheit bei der Datenanonymisierung (bzw. De-Anonymisierung) folgende Fragen zu lösen:

- Welche (rechtlichen) Konsequenzen hat die De-Anonymisierung von Daten?
- Inwiefern entstehen Monitoring-Pflichten zur Überprüfung des (weiterhin) Vorliegens anonymer Daten?
- Wann kann die Kenntnis einer De-Anonymisierung eintreten?
- In welcher Form und in welchem Umfang müssen nach der De-Anonymisierung von Daten datenschutzrechtliche Pflichten erfüllt werden?

4.6 Fazit

Aufgrund des technologischen Fortschritts und der damit verbundenen Verarbeitung immer größer werdender Datenmengen erlangen die Datenverarbeitung in großen Datenverarbeitungssystemen immer größere Bedeutung. Eine (umfangreiche) personenbezogene Datenverarbeitung kann mit erhöhten Risiken für die Rechte und Freiheiten betroffener Personen einhergehen. Die Anonymisierung personenbezogener Daten kann dieser Gefahr – sofern der Verarbeitungskontext eine Verarbeitung rein anonymer Daten zulässt – entgegenwirken.

Die Anonymisierung personenbezogener Daten birgt für Organisationen jedoch ein großes Maß an Rechtsunsicherheit, da es Organisationen häufig schwer fällt einzuschätzen, wann Daten aus rechtlicher Sicht anonym sind. Dies begründet sich nicht zuletzt durch die teils unspezifischen Anforderungen an die Anonymisierung nach der DSGVO. Vor diesem Hintergrund bedarf es klarer Grenzen zwischen personenbezogenen und anonymen Daten sowie Regelungen darüber, wann personenbezogene Daten so hinreichend anonymisiert sind, dass eine Identifizierung der betroffenen Personen nicht mehr möglich ist. Gleichzeitig muss aber auch geklärt werden, welche weiteren Voraussetzungen erfüllt sein müssen, um Daten rechtssicher als „anonym" einstufen zu können. Ziel muss es sein, einerseits zum Schutz betroffener Personen unzureichende Anonymisierungen zu vermeiden sowie andererseits zum Schutz von Organisationen Rechtsunsicherheit im Zusammenhang mit der Anonymisierung zu minimieren.

Fraglich bleibt insofern, wie der Zustand der Anonymität sowie Risiken der De-Anonymisierung gesichert und überwacht werden können. Ein möglicher Weg, um Organisationen in die Lage zu versetzen, die Anonymität eines Datensatzes zu bewerten sowie die Anonymität von Daten zu überwachen, ist die Entwicklung von Metriken und die Verwendung eines Metriken-Systems zur Verifizierung des Anonymitätsgrades eines Datensatzes. Metriken sind mathematische Funktionen, die es erlauben, die Einhaltung von Anforderungen systematisch zu messen und den Ist-Zustand mit dem Soll-Zustand zu vergleichen.[221] Mithilfe von Metriken könnte also auch beurteilt und überprüft werden, ob die Anforderungen an die Anonymität von einem anonymisierten Datensatz erfüllt werden (ggf. in Abhängigkeit vom Anwendungsfall) und ob anonymisierte Daten über die Zeit anonym bleiben.

Die Entwicklung und der Einsatz eines metrischen Systems zur Bewertung und Überprüfung der Anonymität erfordert jedoch zunächst die Lösung der in diesem Kapitel genannten Fragen. Daher müssen die Anonymität und ihre Anforderungen aus einer interdisziplinären Sicht unter Berücksichtigung von Recht und Technologie analysiert werden. Auf der Grundlage der Ergebnisse dieser Analyse sollen in Zukunft die rechtlichen Anforderungen an die Anonymität konkretisiert und überprüfbare Anforderungen und Maßnahmen zur Erreichung von Anonymität abgeleitet werden. Diese überprüfbaren Maßnahmen dienen

[221] Ammann/Sowa, DuD 2012, 247 (247 f.).

dann als Grundlage für eine Reihe von Metriken, anhand derer die Anonymität bewertet, überprüft und überwacht werden kann.

5 IT-Sicherheit in Big-Data-Systemen

In diesem Kapitel soll ausgehend von den Schutzgegenständen zunächst ein Blick auf die spezifischen Gefahren für die datenverarbeitenden IT-Systeme in Big-Data-Umgebungen geworfen werden. Sodann wird der Rechtsrahmen für die IT-Sicherheit in diesen Umgebungen untersucht, und zwar sowohl hinsichtlich der Schutzinstrumente des Datenschutz- als auch des IT-Sicherheitsrechts. Dies geht insbesondere der Frage nach, ob die geltenden rechtlichen Vorgaben und Konkretisierungshilfen für Big-Data-Systeme zum einen normativ angemessen, zum anderen für die Praxis hinreichend handhabbar sind. Auf dieser Basis wird sodann verbleibender Untersuchungs- und Entwicklungsbedarf diskutiert.

5.1 Gefahren für die IT-Systeme der Datenverarbeitung

Im Vordergrund der nachfolgenden Untersuchung steht der Schutz der involvierten IT-Systeme und -Prozesse unter besonderer Berücksichtigung der in ihnen verarbeiteten personenbezogenen Daten. Der Schutz der IT-Systeme ist eine gesamtgesellschaftliche, aber besonders durch den Staat zu gewährleistende Aufgabe,[222] da letzteren eine Schutzpflicht für die Grundrechte und insbesondere für das Grundrecht auf informationelle Selbstbestimmung trifft.[223] Zur Betrachtung der Risiken für personenbezogene Daten durch mangelhaft abgesicherte IT-Infrastrukturen und unsichere Datenverarbeitungssysteme in Big-Data-Anwendungen kann, nach der Erörterung der Schutzgegenstände, die Darstellung übergreifender Gefahren innerhalb der Anwendungsphasen – Beschaffung, Speicherung und Auswertung der Daten[224] – erfolgen.

5.1.1 Schutzgegenstände

Im Vordergrund der Untersuchung steht der Schutz der personenbezogenen Daten in Big-Data-Anwendungen und ihnen zugrundliegenden IT-Systemen. Eine begriffliche Einordnung des Untersuchungsgegenstandes oszilliert zwischen den Begriffen des Datenschutzes, der IT-Sicherheit und der Datensicherheit. Ohne die bislang nicht abschließenden begrifflichen Abgrenzungen zu vertiefen, nimmt der Datenschutz „bei der Auswahl geeigneter technischer und organisatorischer Maßnahmen die Perspektive des Betroffenen und dessen Grundrechtsausübung ein", während die IT-Sicherheit „vorrangig die Informationssicherheit im Blickfeld [hat] und […] die datenverarbeitende Institution schützen"[225] soll. Während der Datenschutz nach Art. 1 Abs. 2 DSGVO auf den Schutz von Grundrechten und Grundfreiheiten natürlicher Personen ausgerichtet ist und seine Gewährleistungsziele erst mithilfe des SDMs (s.u.) konkretisiert werden, schützt die IT-Sicherheit, wie in § 2 Abs. 2 BSIG aufgezählt, dass

- keine unbefugte Person Zugriff auf die IT-Systeme und Daten erlangt (Vertraulichkeit),
- keine Veränderbarkeit der Systeme oder Daten durch unberechtigte Dritte geschieht (Integrität),

[222] Cybersicherheitsstrategie 2021, S. 18 ff.
[223] Demgegenüber ist die Schutzpflicht für das Grundrecht auf Gewährleistung der Integrität und Vertraulichkeit informationstechnischer Systeme, welches das BVerfG im Urt. v. 27. 2. 2008, 1 BvR 370/07 u. 1 BvR 595/07, NJW 2008, 822, entwickelt und für das es die Schutzpflicht in BVerfGE 158, 170 näher begründet hat, bei Big-Data-Systemen typischerweise nicht einschlägig, da es nicht um selbstgenutzte IT-Systeme der Grundrechtsträger geht.
[224] Unterscheidung der drei Phasen für den Datenschutz, *Hackenberg* in *Hoeren/Sieber/Holznagel*, Multimediarecht, Teil 15.2; auch *Fang/Wen/Zheng/Zhou*, IETE Technical Review 2016, S. 3.
[225] *BSI*, IT-Grundschutz-Kompendium, 2023, CON.2, S. 2.

- jederzeitiger Zugriff auf die IT-Systeme und deren Nutzung vorhanden ist, sobald es erforderlich ist (Verfügbarkeit).

Die Schutzziele der IT-Sicherheit weisen Überlappungsbereiche mit denen des Datenschutzes auf. Ohne eine umfassende IT-Sicherheit kann kein Datenschutz gewährleistet werden.[226] Ungeachtet der Abgrenzungsschwierigkeiten[227] zielt der Datenschutz in seinen Anforderungen zum Schutz der Sicherheit der Verarbeitung auf den Schutz der Verarbeitungsprozesse von personenbezogen Daten ab, während die IT-Sicherheit im engeren Sinne sich unmittelbar auf die IT-Systeme bezieht und lediglich die Schutzgüter der IT-Sicherheit in den Fokus nimmt.[228] Abgesehen von den unterschiedlichen Perspektiven sind zumindest die Schutzziele der IT-Sicherheit in denen des Datenschutzes enthalten.

5.1.2 Herausforderungen und Risiken

Die wesentliche Herausforderung in der Sicherheit von Big-Data-Anwendungen (auch und insbesondere bei der Verarbeitung personenbezogener Daten) findet sich in der Integration heterogener Zugriffs- und Verantwortungsbereiche, die in den Händen unterschiedlicher Akteure liegen.[229] Damit sind die Gefahren für die Sicherheit der Big-Data-Systeme nicht allein technischer Natur, sondern ergeben sich auch aus der notwendigen Integration der Nutzer und Anwender, die stets berücksichtigt werden müssen. Ferner sind die Risiken für den Schutz der Informationssysteme und für die personenbezogenen Daten im Einzelfall nicht immer deutlich zu trennen. So können Gefahren für die IT-Systeme zugleich zu Gefahren für den Datenschutz werden.

So ist es nicht ungewöhnlich, dass bis zu einhundert Hersteller an einer IT-Infrastruktur und der mit ihr verbundenen Big-Data-Anwendung eines Unternehmens beteiligt sind.[230] Hierdurch steigt die Gefahr, dass Täter bereits in der Lieferkette ansetzen und scheinbar unbedeutende Teilkomponenten infiltrieren.[231] Zu den Risiken zählen technische Sicherheitslücken in den zugelieferten Komponenten sowie technische Angriffe auf Zulieferer mit Auswirkungen auf das Endprodukt bzw. den Endproduktthersteller. Teilweise wird auch das Personal der Zulieferer selbst nicht so intensiv überprüft oder geschult wie das Personal des Endproduktherstellers.[232] Aufgrund der Unterschiedlichkeit der verwendeten Produkte und der vielzähligen Akteure innerhalb von Big-Data-Anwendungen können sich die bereits bestehenden Probleme in der Lieferkette noch potenzieren.[233] Auch sind intransparente Verarbeitungsketten zu den grundsätzlichen Problemen der immer stärkeren Vernetzung von Big-Data-Systemen zu zählen.

Übergreifend ist zu konstatieren, dass die notwendige Interoperabilität zwischen den verschiedenen Komponenten und Teilanwendungen für eine erhöhte Angreifbarkeit der IT-Systeme sorgt; die Adressierung dieser Herausforderung erlangt in den dezentral eingesetzten Big-Data-Anwendungen besondere Priorität.[234] Ein weiteres Problem ist die Verwendung unterschiedlicher Entwicklungsmodelle für Software, namentlich Open- und Closed-Source-Software, die nicht immer miteinander harmonieren, wodurch sich erhebliche Sicherheitslücken in den Big-Data-Anwendungen ergeben können.[235]

---

[226] *Hornung/Schallbruch* in *Hornung/Schallbruch*: IT-Sicherheitsrecht, § 1 Rdnr. 17.
[227] *Jandt* in *Hornung/Schallbruc*h: IT-Sicherheitsrecht, § 17 Rdnr. 8.
[228] *Jandt* in *Hornung/Schallbruch*: IT-Sicherheitsrecht, § 17 Rdnr. 8 f.
[229] *ENISA*, Big Data Security, S. 4.
[230] *Damm/Fischer*, DuD 2019, 418 (419).
[231] Zu den Lieferkettenangriffen auch *BSI*, BSI-Lagebericht 2022, S. 67, 74.
[232] *BREDEX GMBH*, https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Veranstaltungen/Grundschutz/2GS_Tag_2023/Livehacking_Angriff_ueber_die_Lieferkette.pdf?__blob=publicationFile&v=3.
[233] *ENISA*, Big Data Security, S. 18, 21.
[234] *ENISA*, Big Data Security, S. 15.
[235] *ENISA*, Big Data Security, S. 16.

Darüber hinaus führt die unternehmensübergreifende Vernetzung der IT-Systeme dazu, dass eine Vielzahl von Daten zwischen den früher getrennten IT-Systemen übertragen wird. Dies schließt oftmals die Bestände an personenbezogenen Kunden- und Beschäftigtendaten ein, kann im Falle eines Angriffs aber auch Schadsoftware betreffen. Damit werden Verantwortungs- und Kontrollbereich für die Sicherheit der Daten immer stärker entgrenzt.[236] Auch müssen IT-Sicherheitssysteme mit den hohen Geschwindigkeiten der Datenauswertung, aber auch -beschaffung zurechtkommen, die häufig in Echtzeit geschieht. Entsprechend schnell müssen auch die IT-Systeme einer regelmäßigen Überwachung unterworfen werden, was für die heutigen Systeme zur Angriffserkennung regelmäßig Schwierigkeiten mit sich bringt.[237] Daran anknüpfend stehen auch die zeitaufwendige Sicherheitsmaßnahme der Verschlüsselung und die in Echtzeit agierenden Big-Data-Anwendungen im Konflikt. So kann eine ständige Ver- und Entschlüsselung der differenten Prozesse zu Effizienzeinbußen des Big-Data-Systems führen, was zur Folge hat, dass Verschlüsselungsverfahren in nur wenigen Teilbereichen implementiert werden und dadurch der Datenschutz gefährdet werden kann.

Die Infiltrierung von schädlichen Daten über die Operational-Technology (OT)-Systeme in die Big-Data-Anwendungen gestaltet sich als besonderes Problem in der Phase der Datenbeschaffung. Operational Technology erfasst Soft- und Hardware zur Überwachung und Steuerung v.a. industrieller Anlagen, aber auch anderer Systeme v.a. des Internet of Things. Sie wächst in Zeiten von cyberphysischen Systemen mit der IT zusammen. So können die Täter gezielt schädliche Daten in die vielschichtigen IT-Systeme einschleusen, was oftmals über die OT-Systeme und die mit ihnen verbundenen bzw. durch sie gesteuerten cyberphysischen Systeme geschieht.[238]

In der Speicherungsphase ist zuvorderst anzuführen, dass die Speicherung großer Datenmengen aus einer Vielzahl verschiedener Quellen bereits für sich ein erhöhtes Gefahrenpotenzial mit sich bringt und ein attraktives Ziel für Täter schafft, da bei einem Cyberangriff in kurzer Zeit sehr große Datenbestände und Auswertungsergebnisse erlangt werden können. Je nach Art der Daten kann dies für Angreifer finanziell unmittelbar attraktiv sein (z.B. bei Zahlungsinformationen) oder etwa der Bildung oder Anreicherung von Profilen über eine Vielzahl von Personen dienen. Die Speicherung der gesammelten und ausgewerteten Daten wird mittlerweile durch große Cloud-Infrastrukturen im In- und Ausland ermöglicht. Damit wird die Verantwortlichkeit[239] für die Sicherheit der Cloudumgebung und der Daten – regelmäßig über einen Vertrag – auf Dritte verlagert, die unter Umständen einem anderen Rechtskreis unterliegen,[240] auch wenn sie vertraglich zur Einhaltung technischer und organisatorischer Maßnahmen verpflichtet werden. Gleichzeitig vergrößern sich die notwendigen IT-Strukturen, die für die Übertragung und Speicherung notwendig sind. Damit nehmen die Angriffsvektoren und die Fragilität der IT-Systeme insgesamt zu.

Gleichzeitig werden im Speicherungsprozess verschiedene Datentypen zusammengeführt, die jeweils unterschiedliche Anforderungen an das Schutzbedürfnis der IT-Systeme mitbringen. So ist beispielsweise im Kontext des Datenschutzes zwischen einfachen und sensiblen personenbezogenen Daten zu unterscheiden, wobei letztere eines höheren Schutzniveaus bedürfen. Allerdings kann diese Abstufung in Big-Data-Systemen technisch

[236] *ENISA,* Big Data Threat Landscape and Good Practice Guide, S. 31.
[237] *Von Faber/Kohler*, DuD 2019, 434 (437); *Kohpeiß/Schaller,* CR 2023, 589; auch *Fang/Wen/Zheng/Zhou*, IETE Technical Review 2016, S. 6.
[238] *Börner/Koepke/Zenger* in *Vogt/Hennies/Endreß/Peters*, Wirtschaftsschutz in der Praxis, S. 249 f., 252.
[239] Im Kontext der DSGVO handelt es sich dabei häufig um eine Auftragsverarbeitung i.S.v. Art. 28 i.V.m. 4 Nr. 8 DSGVO; zu den Verantwortlichkeiten etwa *Ingold* in *Sydow/Marsch,* DSGVO/BDSG, Art. 28 Rdnr. 11 m.w.N.
[240] *Völker/Schnatz/Breyer*, MMR 2022, 427 (431); auch *Rafiq/Awan/Yasin/Nobanee/Zain/Bahaj*, Sage journals 2022, 1 (14).

schwierig sein,[241] und aus Anwendersicht mag bei einem kleinen Anteil besonders schützenswerter Daten eine Orientierung am höchsten Schutzbedarf unverhältnismäßig erscheinen. Deshalb besteht besonders in den heterogenen Big-Data-Anwendungen die Gefahr, dass sich am Schutzniveau der weniger wichtigen Daten orientiert wird.

In der Phase der Auswertung muss insbesondere der Aspekt berücksichtig werden, dass Big-Data-Systeme oftmals mit Algorithmen künstlicher Intelligenz operieren. Dies betrifft sowohl die Analyse der Datensammlungen als auch die anschließende Selbstoptimierung des KI-Systems. Systeme künstlicher Intelligenz können wichtige Beiträge zur Stärkung der IT-Sicherheit leisten, aufgrund ihrer komplexen technischen Grundlagen aber auch erhebliche IT-Sicherheitsprobleme aufweisen.[242] Verschaffen Angreifer sich über ein ungesichertes IT-System Zugang zu den Trainingsdaten, können Manipulationen der Daten zu erheblichen und zum Teil grundrechtsgefährdenden Ergebnisveränderungen führen.[243] Konkret können Sicherheitslücken in der Integrität der Lerndaten und des Lernverfahrens sowie in der Kennzeichnung und Einstufung der notwendigen Parameter bestehen.[244]

5.2 Regulierungsrahmen

Ein einheitlicher Regulierungsrahmen für die Sicherheit von IT-Systemen[245] fehlt bislang ebenso wie spezifische IT-sicherheitsrechtliche Maßgaben für Big-Data-Anwendungen und die in ihnen enthaltenen Daten. Allerdings können sich Vorgaben, die für den Schutz der datenverarbeitenden IT-Systeme in Big-Data-Anwendungen relevant sind, aus IT-sicherheitsrechtlichen Vorgaben des Datenschutzrechts sowie aus allgemeinen sowie sektorspezifischen IT-Sicherheitsregulierung ergeben. Hervorzuheben sind insbesondere die Vorgaben der DSGVO zum Schutz personenbezogener Daten (und damit mittelbar der datenverarbeitenden IT-Systeme) sowie (weithin europarechtlich determinierte) Regelungen im BSIG, welches die größte Vorschriftensammlung des IT-Sicherheitsrechts darstellt und neben weiteren Spezialgesetzen steht.[246] Schließlich bestehen weitere (europäische) Gesetze und Einzelvorschriften, die sich mit Fragen der Datensicherheit im Allgemeinen befassen. Auf EU-Ebene zählen dazu auch die fünf – bereits in Kraft getretenen oder noch in der Entwurfsfassung steckenden – Datenrechtsakte.[247] So enthält beispielsweise der Entwurf für eine KI-Verordnung (dazu auch Kap. 3.5.1) Vorgaben für die Sicherheit von IT-Systemen, die den Schutz von Big-Data-Anwendungen ergänzen werden.[248] Nachfolgend liegt der Fokus auf den Maßgaben des BSIG und der DSGVO, welche die wesentlichen IT-sicherheitsrechtlichen Instrumente zum Schutz der Big-Data-Systeme enthalten.

5.2.1 Systematik

Die IT-sicherheitsrechtlichen Vorgaben des Datenschutzrechts ergeben sich vor allem aus Art. 32 DSGVO, der sich – je nach Verständnis der Abgrenzung – als eine Norm verstehen lässt, die sowohl dem IT-Sicherheits- als auch dem Datenschutzrecht zuzuordnen ist.[249]

---

[241] Für die Klassifizierung von E-Mails für die datenschutzkonforme Löschung und Archivierung s. *Kunz/Waldmann,* INFORMATIK 2022, S. 589 ff.
[242] Dazu *Djeffal*, MMR 2019, 289 (293f.), *Müller-Quade/Meister/Holz/Houdeau/Rieck/Rost/Schauf/Schindler,* Künstliche Intelligenz und IT-Sicherheit – Bestandsaufnahme und Lösungsansätze, 2019; s.a. die Beiträge in *Ebers/Steinrötter*, Künstliche Intelligenz und smarte Robotik im IT-Sicherheitsrecht, 2021.
[243] *Hackenberg* in *Hoeren/Sieber/Holznagel*: Multimediarecht, Teil 15.2 Rdnr. 61 f.
[244] *Von Faber/Kohler,* DuD 2019, 434 (437).
[245] *Conrad* in *Auer-Reinsdorff/Conrad*: IT- und Datenschutzrecht, § 33 Rdnr. 9.
[246] Sowohl das BSIG als auch die IT-sicherheitsrechtlichen Vorgaben der DSGVO können als allgemeines IT-Sicherheitsrecht umschrieben werden *Raabe/Schallbruch/Steinbrück*, CR 2018, 706 (707).
[247] *Debus i*n Gersdorf/Paal, BeckOK InfoMedienR, § 1 Rdnr. 10; zu den Auswirkungen der EU-Datenstrategie auf den Datenschutz s. *Kipker,* ZD-Aktuell 2022, 04465; zu einem Quervergleich mit Blick auf die Vorschriften mit Bezug zum Datenschutz, zur Interoperabilität und zur sonstigen technischen Gestaltung (einschließlich IT-Sicherheit) s. *Pfeiffer/Helmke*, ZD-Aktuell 2023, 01162.
[248] Vgl. Art. 15 KI-VO-E, COM/2021/206 final, 21.04.2021.
[249] S. *Hornung/Schallbruch* in *Hornung/Schallbruch*: IT-Sicherheitsrecht, § 1 Rdnr. 17; enger *Jandt* in *Hornung/Schallbruch*: IT-Sicherheitsrecht, § 17 Rdnr. 33 ff.

Unabhängig von der Größe oder Bedeutung der datenverarbeitenden Einrichtung, die Big-Data-Systeme anwendet, hat der Verantwortliche für den Schutz der Verarbeitung personenbezogener Daten nach Art. 32 DSGVO technisch-organisatorische Maßnahmen zu implementieren. Voraussetzung ist die Eröffnung des sachlichen Anwendungsbereichs der DSGVO, der nach Art. 2 Abs. 1 iVm. Art. 4 Nr. 1, 2 DSGVO die Verarbeitung personenbezogener Daten voraussetzt. Es bedarf daher Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen[250] und die innerhalb eines Vorgangs - wie Erheben, Erfassen, Speichern, Übermitteln oder Löschen - verarbeitet werden.[251]

Hingegen beziehen sich die Vorgaben des BSIG und auch die einzelnen sektorspezifischen IT-Sicherheitsbestimmungen nach § 2 Abs. 2 S. 4 BSIG auf „die Einhaltung bestimmter Sicherheitsstandards, die die Verfügbarkeit, Integrität oder Vertraulichkeit von Informationen betreffen, durch Sicherheitsvorkehrungen in informationstechnischen Systemen, Komponenten oder Prozessen oder bei der Anwendung von informationstechnischen Systemen, Komponenten oder Prozessen".[252] Die – maßgeblich im BSIG – abgebildeten Adressaten des IT-Sicherheitsrechts beschränken sich auf wenige Gruppen von Betreibern, die allerdings über die letzten acht Jahre sukzessive erweitert wurden. Zunächst adressierte der Gesetzgeber in § 8a BSIG Betreiber Kritischer Infrastrukturen, die nach der BSI-KritisV bestimmt werden und entsprechende Dienstleistungen von entsprechendem Umfang in der allgemeinen Daseinsvorsorge erbringen (vgl. auch § 2 Abs. 10 BSIG). Daneben wurden im Bereich der Kritischen Infrastrukturen vereinzelt auch Unternehmen reguliert, die dem KRITIS-Bereich zugehören, jedoch nicht die notwendigen Schwellenwerte der BSI-KritisV erreichen, wie einzelne Energienetzbetreiber, Telekommunikationsanbieter oder Finanzinstitute. Seit 2017 werden auch Anbieter digitaler Dienste nach § 8c BSIG den IT-Sicherheitsbestimmungen unterworfen. Mit dem IT-Sicherheitsgesetz 2.0 erfolgte in § 8f BSIG eine Ausweitung auf Schlüsselunternehmen (sog. Unternehmen im besonderen öffentlichen Interesse), die eine große Bedeutung in Bezug auf die IT-Sicherheit in Deutschland haben.[253]

Etliche dieser Adressaten sind Unternehmen, die in – tlw. sehr – großem Stil personenbezogene Daten verarbeiten. In diesem Sinne schafft das IT-Sicherheitsrecht, sofern es nach dem Vorstehenden anwendbar ist, eine Regulierung, die zwar im Sinne von § 2 Abs. 2 S. 4 BSIG nicht die datenschutzrechtlich betroffene Person zum Ausgangspunkt nimmt, aber zumindest mittelbar auch dem Schutz personenbezogener Daten dient. Dies ist besonders deutlich bei den Anbietern digitaler Dienste im Sinne von § 2 Abs. 11 BSIG. Seit dem Jahre 2017 enthält das Gesetz auch Vorgaben für diese Dienste, die Online-Marktplätze, Online-Suchmaschinen und Cloud-Computing-Dienste umfassen. Damit werden einige Big-Data-Anwendungen, allerdings bei weitem nicht alle, vom IT-Sicherheitsrecht erfasst.

5.2.2 Schutzinstrumente der DSGVO

Nachfolgend sollen die Chancen und Herausforderungen des Art. 32 DSGVO für den Schutz der spezifischen IT-Systeme beleuchtet werden. Besonders das Standard-Datenschutzmodell (SDM) und die selbstregulatorischen Ansätze der DSGVO bilden wichtige Anknüpfungspunkte, um die Anforderungen des Art. 32 DSGVO zu konkretisieren.

5.2.2.1 Technisch-organisatorische Maßnahmen nach Art. 32 DSGVO

Nach Art. 32 DSGVO haben sämtliche Verantwortliche und Auftragsverarbeiter angemessene technisch-organisatorische Maßnahmen zum Schutz der personenbezogenen Daten

[250] Zur Frage des Personenbezugs näher *Schild* in *Wolff/Brink/v. Ungern-Sternberg*, BeckOK DatenschutzR, Art. 4 Rdnr. 3.
[251] *Schild* in *Wolff/Brink/v. Ungern-Sternberg*, BeckOK DatenschutzR, Art. 4 Rdnr. 29.
[252] *Hornung/Schallbruch* in *Hornung/Schallbruch*: IT-Sicherheitsrecht, § 1 Rdnr. 10.
[253] Zum IT-Sicherheitsgesetz 2.0 s. *Hornung,* NJW 2021, 1985 (1988).

zu implementieren und diese auch zu dokumentieren.[254] Die Maßnahmen sind vor allem unter Berücksichtigung des Stands der Technik und der Implementierungskosten, aber auch der „Art, des Umfangs, der Umstände und der Zwecke der Verarbeitung sowie der unterschiedlichen Eintrittswahrscheinlichkeit und Schwere des Risikos für die Rechte und Freiheiten natürlicher Personen" auszuwählen und umzusetzen.[255] Zur Gewährleistung der Sicherheit normiert Art. 32 Abs. 1 DSGVO bestimmte Standardmaßnahmen und gewünschte Eigenschaften, nämlich

- die Pseudonymisierung und Verschlüsselung personenbezogener Daten (lit. a),
- die Fähigkeit, die Vertraulichkeit, Integrität, Verfügbarkeit und Belastbarkeit der Systeme und Dienste im Zusammenhang mit der Verarbeitung auf Dauer sicherzustellen (lit. b);
- die Fähigkeit, Verfügbarkeit der personenbezogenen Daten und den Zugang zu ihnen bei einem physischen oder technischen Zwischenfall rasch wiederherzustellen (lit. c)
- Verfahren zur regelmäßigen Überprüfung, Bewertung und Evaluierung der Wirksamkeit der technischen und organisatorischen Maßnahmen zur Gewährleistung der Sicherheit der Verarbeitung (lit. d).

Diese Vorgaben und die insoweit genannten Maßnahmen und Schutzziele sind nicht überschneidungsfrei, geben aber einen verbindlichen Rahmen für die durch Verantwortliche und Auftragsverarbeiter zu ergreifenden Maßnahmen. Besonders hervorzuheben – auch mit Blick auf die Schnittstelle zwischen dem Datenschutz- und dem IT-Sicherheitsrecht – ist Art. 32 Abs. 1 lit. b DSGVO, der auch die IT-Schutzziele konkret benennt.

Problematisch ist jedoch, dass die genannten Vorgaben abstrakt formuliert sind und in der Anwendung des risikobasierten Ansatzes (Art. 32 Abs. 2 DSGVO)[256] wenig Rechtssicherheit zu finden ist.

5.2.2.2 Operationalisierung durch das Standarddatenschutz-Modell (SDM)

Die Praxis orientiert sich deshalb an verschiedenen Maßnahmenkatalogen wie dem IT-Grundschutz des BSI und den – zwar im Verhältnis zu Art. 32 DSGVO deutlich konkreteren, allerdings immer noch relativ abstrakten und technologieübergreifenden – Katalogen der ISO/IEC 27001-Reihe.[257]

Allerdings legten diese Instrumente in der Vergangenheit den Fokus nicht auf den Grundrechtsschutz, den das Datenschutzrecht ausweislich Art. 1 Abs. 1 und Abs. 2 DSGVO bezweckt. Daher werden die überwiegend abstrakt formulierten Anforderungen aus Art. 32 DSGVO, genauso wie die übrigen Vorgaben aus der DSGVO, durch verschiedene datenschutzorientierte Konkretisierungen in handhabbarere Formen überführt.

Zum einen bewegt sich die eher klassische IT-Sicherheitsstandardisierung in dieser Richtung, nämlich mit der ISO/IEC 27701 (Privacy Information Management System). Diese ist als globale Norm nicht auf die Anwendung der DSGVO beschränkt, sondern ermöglicht einer Organisation die Einführung und regelmäßige Überprüfung eines Managements des internen Datenschutzes in Übereinstimmung mit (europäischen oder anderen) Datenschutzvorgaben, mit dem Ziel, das Management kontinuierlich an neue Anforderungen anzupassen.

---

[254] *Deusch/Eggendorfer* in *Taeger/Pohle*: Computerrecht, 50.1 Rdnr. 310.
[255] *Paulus* in *Wolff/Brink/v. Ungern-Sternberg*, BeckOK DatenschutzR, Art. 32 Rdnr. 7.
[256] S. zu diesem für das Beispiel des Art. 32 DSGVO *Hansen* in *Simitis/Hornung/Spiecker gen. Döhmann*, DSGVO/BDSG, Art. 32 Rdnr. 58 ff.
[257] S. aus datenschutzrechtlicher Sicht *Hansen* in *Simitis/Hornung/Spiecker gen. Döhmann*, DSGVO/BDSG, Art. 32 Rdnr. 78.

Zum anderen nimmt das Standard-Datenschutzmodell (SDM) das Datenschutzrecht zum Ausgangspunkt und transformiert dessen generische Anforderungen in sieben Gewährleistungsziele, um die rechtlichen Vorgaben der DSGVO operabel zu machen.[258] Anders als die ISO/IEC 27701 konzentriert sich das SDM also auf Prinzipien der DSGVO und leitet aus ihnen die Gewährleistungsziele Datenminimierung, Verfügbarkeit, Integrität, Vertraulichkeit, Transparenz, Nichtverkettung und Intervenierbarkeit ab, wodurch mit Blick auf die sich teilweise überschneidenden Schutzziele der IT-Sicherheit auch die Schnittstelle zwischen Datenschutz und IT-Sicherheit deutlich wird.

Das SDM dient der Operationalisierung der technisch-organisatorischen Maßnahmen aus Art. 32 DSGVO und unterstützt bei der Bewertung und Auswahl angemessener Maßnahmen. Nach einer Unterteilung der Datenverarbeitung in die Bereiche Daten, Systeme und Prozesse erfolgt eine auf den Schutz der betroffenen Personen ausgerichtete Schutzbedarfsanalyse. Dadurch werden die Vorgaben aus der DSGVO durch verschiedene datenschutzorientierte Konkretisierungen in handhabbarere Formen überführt, und es wird eine Transformationshilfe zwischen Recht und Technik geschaffen.[259] Als solche wird das SDM inzwischen auch durch den BSI-Grundschutzkatalog anerkannt.[260] Die Schwäche des SDM liegt jedoch darin, dass die Ausgestaltung der Verhältnismäßigkeit beim Verantwortlichen verbleibt und dieser für Big-Data-Systeme weder konkrete Kriterien für die Abwägung der Gefahren mit den wirtschaftlichen Interessen noch konkrete Maßnahmenvorschläge erhält. Mit anderen Worten ist es zwar grundsätzlich methodisch möglich, Vorgaben für derartige Systeme und die erheblichen mit ihnen verbunden Risiken abzuleiten; die Konkretisierungsleistung, den der risikobasierte Ansatz des Art. 32 DSGVO dem Verantwortlichen auferlegt, muss dieser jedoch selbst erbringen.

Daneben ist Teil des SDM, dass „vorrangig die Gewährleistungsziele mit Datenschutzbezug" betrachtet werden (vgl. auch Art. 4 Nr. 2 DSGVO, auch Art. 32 Abs. 1 lit. b DSGVO)[261] und somit keine Anforderungen mit einbezogen werden, die über das Datenschutzrecht hinausgehen. Damit bleiben die zugrundeliegenden IT-Systeme, die nicht unmittelbar der Datenverarbeitung dienen, für die Gesamtsicherheit der Organisation aber relevant sein können, mehrheitlich unberücksichtigt.

5.2.2.3 Zertifizierung und Verhaltensregeln in der DSGVO

Bedeutende Instrumente zur Gewährleistung technisch-organisatorischer Maßnahmen und für die rechtskonforme Gestaltung von Big-Data-Anwendungen könnten Verhaltensregeln und Zertifizierungen nach Art. 40 und 42 DSGVO sein (so auch Art. 32 Abs. 3 DSGVO).

Nach Art. 40 Abs. 1 DSGVO fördern die „Mitgliedstaaten, die Aufsichtsbehörden, der Ausschuss und die Kommission […] die Ausarbeitung von Verhaltensregeln […]", wozu grundsätzlich auch solche über „Maßnahmen und Verfahren gemäß den Artikeln 24 und 25 und die Maßnahmen für die Sicherheit der Verarbeitung gemäß Artikel 32" gehören. Mit den Verhaltensregeln wird nach ErwG. 13 den „Bedürfnisse[n] von Kleinstunternehmen sowie kleinen und mittleren Unternehmen"[262] entgegengekommen.[263] Grundsätzlich wären Verhaltensregeln für eine Operationalisierung der technisch-organisatorischen Vorgaben nach Art. 32 DSGVO unter Berücksichtigung der Big-Data-Anwendungen geeig-

[258] *DSK,* Standard-Datenschutzmodell, Version 3.0, 2022; näher *Rost*, PinG 2023, 94; ausführlich *Rost*, Das Standard-Datenschutzmodell (SDM), 2022.
[259] *DSK*, Standard-Datenschutzmodell, Version 3.0, 2022, S. 7.
[260] S. *BSI,* BSI-Grundschutzkatalog, Baustein CON.2, Edition 2023; das SDM entspricht auch den Empfehlungen des IT-Planungsrates, s. Beschluss vom 25.03.2020.
[261] *DSK*, Standard-Datenschutzmodell, Version 3.0, 2022, S. 65 f.
[262] Empfehlung der Kommission vom 6. Mai 2003 betreffend die Definition der Kleinstunternehmen sowie der kleinen und mittleren Unternehmen (C (2003) 1422) (ABl. L 124 vom 20.5.2003, S. 36).
[263] *Roßnagel* in *Simitis/Hornung/Spiecker gen. Döhmann*, DSGVO/BDSG, Art. 40 Rdnr. 24.

net. Jedoch wird das Instrument der Verhaltensregeln auf nationaler als auch auf europäischer Ebene bisher kaum genutzt[264]. An entsprechenden Vorschlägen für die IT-Sicherheit in Big-Data-Anwendungen fehlt es demensprechend bislang.

Eine weitere Chance zur Absicherung der IT-Systeme von Big-Data-Anwendungen könnte in der freiwilligen Zertifizierung nach Art. 42 DSGVO liegen, die sich auf die Prüfung und Bewertung von Verarbeitungsvorgängen bezieht (vgl. Abs. 1 S. 1).[265] Ähnlich wie bei Selbstregulierung durch Verhaltenspflichten haben auch hier die Aufsichtsbehörden sowie die Mitgliedsstaaten eine in Art. 42 Abs. 1 DSGVO normierte Förderpflicht.[266] Sinnvoll wäre die Bewertung eines einheitlichen und abgrenzbaren Verarbeitungssystems, in dem die Risiken der Verarbeitung ganzheitlich beurteilt werden können.[267] Nachdem die Zertifizierungskriterien durch die zuständige Stelle genehmigt wurden, können staatliche oder private akkreditierte Stellen eine Zertifizierung vornehmen. Anders als bei den Verhaltensregeln steht die Zertifizierung derzeit an der Schwelle zu einem breiteren Einsatz. Allerdings geht der Trend bisher dahin, generische Kriterienkataloge zu erarbeiten, bei denen die erforderlichen Konkretisierungen im Rahmen der einzelnen Zertifizierung durch die Zertifizierungsstelle erbracht werden müssen.[268] Ein spezifisch auf Big Data und/oder KI-basierte Datenverarbeitungen ausgerichtetes Zertifizierungsprogramm, das spezifische Risiken berücksichtigt und auf diese angepasste Prüfkriterien formuliert, könnte ein echter Gewinn für die Verpflichteten sein, ist bisher aber nicht in Sicht.

5.2.3 Schutzinstrumente des BSIG

Zum Instrumentenkasten des IT-Sicherheitsrechts gehören u.a. Meldepflichten der Unternehmen, Nachweise, Dokumentierungen, Tiefenprüfungen, Zertifizierungen, Warnungen sowie Implementierungspflichten für spezifische Sicherheitskomponenten, die sich jedoch je nach Adressaten leicht unterscheiden können. Ähnliche IT-Sicherheitspflichten weisen die sektorspezifischen Vorschriften auf, die zu den Vorgaben des BSIG hinzutreten oder sie ersetzen und sich im Atom-, Energie-, Finanz- und Versicherungs-, Gesundheits- und Telekommunikationssektor finden. Besonders zu berücksichtigen sind nachfolgend die Instrumente, die eine Erweiterung auf die spezifischen Anforderungen der IT-Systeme der Datenverarbeitung im Kontext von Big-Data-Anwendungen zulassen und auch für kleine Unternehmen zugeschnitten werden können.

5.2.3.1 BSI-Grundschutz

Auf Ebene des Soft Law existiert in der IT-Sicherheit als Pendant zum SDM im Datenschutz der sog. BSI-Grundschutz. Er ist ein freiwilliges Instrument, das entweder bei der Zertifizierung oder Normkonkretisierungen – etwa zur Ausfüllung des Rechtsbegriffs „Stand der Technik" – unterstützt.[269] Für sämtliche regulierten sowie nicht regulierten Unternehmen findet der BSI-Grundschutz Anwendung. Mit seinen technischen, personellen und organisatorischen Umsetzungshilfen dient er der umfassenden Sicherheit der informationstechnischen Systeme und bietet den Verantwortlichen vielfältige Best-Practice Verfahren an. Damit bietet das BSI zwar ein umfassendes Basis-Schutzkonzept für alle Unternehmen an, welches sich unabhängig von der Unternehmensgröße und der verwendeten (heterogenen) Technologien anwenden lässt. Jedoch geht es nicht auf die spezifische Gefahrenlage der vielschichten Big-Data-Anwendungen ein.

---

[264] *Jungkind* in *Wolff/Brink/v. Ungern-Sternberg*, BeckOK DatenschutzR, Art. 40 Rdnr. 1.
[265] *Scholz* in *Simitis/Hornung/Spiecker gen. Döhmann*, DSGVO/BDSG, Art. 42 Rdnr. 21; *Hornung/Hartl,* ZD 2014, 219 (224).
[266] *Jungkind* in *Wolff/Brink/v. Ungern-Sternberg*, BeckOK DatenschutzR, Art. 40 Rdnr. 7 sieht ein „aktives Element".
[267] *Scholz* in *Simitis/Hornung/Spiecker gen. Döhmann*, DSGVO/BDSG, Art. 42 Rdnr. 22.
[268] Dies gilt etwa für den luxemburgischen GDPR-CARPA (GDPR-Certified Assurance-Report based Processing Activities; dazu EDSA, Opinion 1/2022; *Helmke/Link/Schild,* DuD 2023, (100) und das European Privacy Seal (EuroPriSe), das sich an Auftragsverarbeiter richtet, s. EDSA, Opinion 28/2022.
[269] *Djeffal*, MMR 2019, 289 (292).

5.2.3.2 Branchenspezifische Standards nach dem BSIG

Betreiber kritischer Infrastrukturen oder ihre Branchenverbände haben nach § 8a Abs. 3 S. 1 BSIG die Möglichkeit, „branchenspezifische Sicherheitsstandards zur Gewährleistung der Anforderungen nach den Absätzen 1 und 1a" vorzuschlagen. Für die übrigen Adressaten in der IT-Sicherheitsregulierung fehlt diese Möglichkeit jedoch bislang und stellt somit kein flexibel einsetzbares und verallgemeinerbares Schutzinstrument dar. Anbieter digitaler Dienste müssen zwar angemessene technisch-organisatorische Maßnahmen umsetzen (§ 8c BSIG), haben jedoch keine Möglichkeit, Standards selbst zu entwerfen. Unternehmen im besonderen öffentlichen Interesse nach § 2 Abs. 14 BSIG wurde die Möglichkeit gewährt, ihre IT-Sicherheitsmaßnahmen durch Selbsterklärungen nachzuweisen (§ 8f Abs. 1 BSIG). Damit enthält das BSIG zwar Ansätze einer regulierten Selbstregulierung, jedoch insoweit beschränkt auf wenige Adressaten.

5.2.3.3 Zertifizierungen und IT-Sicherheitskennzeichen

Indes könnte die Zertifizierung nach §§ 9, 9a BSIG eine entscheidende Rolle für den Schutz von Big-Data-Systemen spielen. Nach § 9 BSIG kann für „bestimmte Produkte oder Leistungen […] beim Bundesamt eine Sicherheits- oder Personenzertifizierung oder eine Zertifizierung als IT-Sicherheitsdienstleister beantragt werden", jedoch nur, wenn „informationstechnische Systeme, Komponenten, Produkte oder Schutzprofile den vom Bundesamt festgelegten Kriterien entsprechen". Konkretisiert wird die Zertifizierung nach § 9 BSIG durch die BSI-Zertifizierungs- und -Anerkennungsverordnung (BSIZertV). Das BSI hat bereits umfassende Listen mit zertifizierten IT-Systemen und Teilkomponenten, die wöchentlich wachsen.[270] Gleichwohl fehlt es bislang an einem umfassenden und zertifizierbaren Schutzkonzept und genehmigten Kriterien des BSI, um Big-Data-Systeme zu erfassen und umfassend bewerten zu können. Für die europäische Cybersicherheitszertifizierung legt § 9a BSIG fest, dass das BSI die Bewertungsstelle nach Art. 58 Absatz 1 der Verordnung (EU) 2019/881 wird. Bislang fehlt es jedoch auch an einer spezifischen Zertifizierungsmöglichkeit von Big-Data-Systemen nach einem europäischen Schema.

Schließlich ist noch an das IT-Sicherheitskennzeichen nach § 9c BSIG zu denken, welches mit dem IT-Sicherheitsgesetz 2.0 eingefügt wurde und für vom Bundesamt festgelegte Produktkategorien eine Konformitätsbewertung und anschließende Kennzeichnung ermöglicht.[271] Es wird jedoch in der Norm bereits explizit darauf hingewiesen, dass nur Schutzanforderungen der IT-Sicherheit, jedoch keine des Datenschutzes beachtet werden.

5.2.4 Zwischenergebnis

Betrachtet man die oben erläuterten Herausforderungen innerhalb von Big-Data-Anwendungen, so ergeben sich vielschichtige Schutzanforderungen, die das Recht bislang unzureichend erfasst. Je nachdem, wo man ansetzt, fällt auf, dass einerseits deutlich zu wenig Adressaten erfasst sind oder die Schutzinstrumente nicht auf die Anforderungen der Big-Data-Systeme abgestimmt sind. Gleichzeitig wird deutlich, dass die für Big-Data-Sicherheit gewinnbringenden Instrumente maßgeblich auf der Eigeninitiative der nichtregulierten Unternehmen beruhen und keine Verbindlichkeit besitzen.

5.3 Quo vadis?

Wie soll es im Schutz um die IT-Sicherheit von Big-Data-Anwendungen weitergehen? Es bestehen verschiedene Möglichkeiten, wie der Gesetzgeber auf die fehlenden konkreten

[270] https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/Zertifizierung-und-Anerkennung/Listen/listen_node.html.
[271] *Hornung,* NJW 2021, 1985 (1989).

Maßgaben zum Schutz von Big-Data-Anwendungen und der in ihnen verarbeiteten personenbezogenen Daten reagieren kann. Zukünftige Rechtsentwicklungen müssen beachten, dass der technisch-organisatorische Datenschutz und die Sicherheit in den IT-Systemen gemeinsam „gedacht" und umgesetzt werden sollten. Schließlich erscheinen ganzheitliche (formelle) gesetzliche Regelungen für die heterogenen Big-Data-Anwendungen kaum realisierbar.[272] Vielmehr können einfachgesetzliche Verbesserungsvorschläge nur in Teilbereichen von Big Data ansetzen. Daneben sollte jedoch auf der Umsetzungsebene verstärkt versucht werden, ganzheitliche Schutzkonzepte zu entwickeln und einzubringen. Nachfolgend sollen erste Lösungsansätze formuliert werden, die als Grundlage für weiteren Forschungsbedarf dienen sollen.

5.3.1 Ausweitung der Adressaten und Verpflichtungen

Eine naheliegende Möglichkeit zur Verbesserung des Schutzes der IT-Systeme wäre eine Erweiterung des Adressatenkreises in der IT-Sicherheitsregulierung. Hierdurch könnte zunächst für sämtliche (auch teilweise ausgelagerte) IT-Systeme eine grundsätzliche Erhöhung des Schutzniveaus bewirkt werden, auf dem weiter aufgebaut werden könnte. Dies entspricht der aktuellen gesetzgeberischen Intention: So werden im Referentenentwurf zur Umsetzung der NIS-2-RL[273] die Adressaten und zugleich ihre Pflichten erheblich erweitert.[274] Danach sollen zukünftig Unternehmen aus 18 Sektoren angemessene technisch-organisatorische IT-Sicherheitsmaßnahmen verpflichtend umsetzen müssen. Sämtliche „mittlere" Unternehmen mit hoher Kritikalität (Anhang 1 der RL) und Unternehmen, die Teil der sonstigen kritischen Infrastrukturen sind (Anhang 2 der RL) sowie weitere Schlüsselunternehmen werden der Regulierung unterfallen. Unbeachtet bleiben damit jedoch weiterhin zahlreiche insbesondere kleine und mittlere Unternehmen mit Big-Data-Anwendungen. Es wäre deshalb näher zu untersuchen, ob Unternehmen, die derartige Anwendungen in Betrieb halten, unabhängig von der Unternehmensgröße reguliert werden sollen

Ein zweiter Ansatz könnte in einer spezifischen IT-sicherheitsbezogenen Regulierung der Verarbeitung von Big Data durch komplexe Algorithmen liegen. Einen solchen Weg geht der Entwurf der Europäischen Kommission für eine KI-VO.[275] Dieser enthält in Art. 15 KI-VO-E Vorgaben für die „Genauigkeit, Robustheit und Cybersicherheit" der Systeme (vgl. auch ErwG. 51). Nach Abs. 3 müssen die Systeme widerstandsfähig gegenüber Fehlern, Störungen oder Unstimmigkeiten sein; Robustheit wird durch technische Redundanz erreicht. Insbesondere schreibt Abs. 4 vor, dass die Systeme widerstandfähig gegen Versuche unbefugter Dritter sein müssen, ihre Verwendung oder Leistung durch Ausnutzung von Systemschwachstellen zu verändern. Es sind angemessene technische Maßnahmen zu ergreifen, die auch KI-spezifische Schwachstellen wie die Manipulation der Trainingsdaten einbeziehen. Darüber hinaus präzisiert der Vorschlag in Art. 15 Abs. 4 KI-VO-E allerdings weder den Inhalt noch das Niveau der IT-Sicherheit.[276] Außerdem finden die Bestimmungen nur auf sog. „Hochrisiko-KI-Systeme" Anwendung. Dazu zählen nach Art. 6 Abs. 1 KI-VO-E KI-Systeme, die unter produktorientierte Harmonisierungsvorschriften der Union fallen, sowie nach Art. 6 Abs. 2 KI-VO-E auch sämtliche KI-Systeme aus Anhang III, die in verschiedenen kritischen Infrastrukturen und weiteren Bereichen wie biometrische Identifizierung, allgemeine und berufliche Bildung, Beschäftigung und Personalmanagement, bestimmte grundlegende private und öffentliche Dienste und Leistungen, Strafverfolgung, Migration, Asyl und Grenzkontrolle sowie Rechtspflege und demokratische Prozesse eingesetzt werden. Damit überlappt dieser Regulierungsansatz mit Teilen der bestehenden Regulierung kritischer Infrastrukturen, geht aber in anderen Bereichen deutlich

[272] S. auch *Hornung* in *Hoffmann-Riem*: Big Data – Regulative Herausforderungen, S. 95.
[273] RL 2022/2555, ABl. L333/80.
[274] Synopse eines Referentenentwurfs, abrufbar unter: https://intrapol.org/2023/05/10/refe-fuer-ein-nis-2-umsetzungs-und-cybersicherheitsstaerkungsgesetz-nis2umsucg/; s. auch *Kipker/Dittrich*, MMR 2023, 481.
[275] Zu weiteren Fragen des Entwurfs und den Auswirkungen auf Schutzrechte s. Kap. 3.5.1.
[276] *Deusch/Eggendorfer* in *Taeger/Pohle*: Computerrecht, 50.1 Rdnr. 280c.

darüber hinaus. Damit werden bedeutende, nämlich grundrechtssensible Teilbereiche von Big-Data-Anwendungen erfasst und verantwortliche Adressaten zu IT-Sicherheitsmaßnahmen angehalten.

5.3.2 Stärkung der Zertifizierungs- und Verhaltensregeln

Die Möglichkeit der Verhaltensregeln und Zertifizierungen aus Art. 40, 42 DSGVO könnte ein wichtiges Instrument zur Etablierung von IT-Sicherheitsstandards sein. Dementsprechend ist die Förderpflicht von Mitgliedstaaten, Aufsichtsbehörden, Europäischem Datenschutzausschuss und Kommission ein Hebel, der ein erhebliches Potenzial verspricht. Dies gilt vor allem, wenn der Blick auf die Gefahren der IT-Sicherheit erweitert werden kann, die gerade aus dem Zusammenspiel der Teilbereiche von großen und heterogenen Datenmengen, Cloud Computing sowie KI entstehen. Eine Herausforderung besteht freilich darin, unabhängige Prüfungen und Zertifizierungen nach der DSGVO mit den anderen, stärker IT-sicherheitstechnisch ausgerichteten Zertifizierungen zu verzahnen. Die Zertifizierungslandschaft wird in Zukunft durch die neuen Vorgaben der KI-Verordnung noch komplexer werden. Eine Gesamt-Zertifizierung, die alle Vorgaben einschließt, ist bisher regulatorisch nicht in Sicht, dürfte eine erhebliche Komplexität aufweisen und würde deshalb einen erheblichen Forschungs- und Vorbereitungsaufwand erzeugen. Solange hierfür keine Pläne absehbar sind, sollte zumindest auf Ebene der Prüfkriterien eine möglichst starke Abstimmung erfolgen.

Daneben bestünde auch die Möglichkeit, ein selbstregulatorisches Instrument der Verhaltensregeln mit einer entsprechenden Förderpflicht des BSI auch im IT-Sicherheitsrecht zu verankern. Insbesondere KMUs hätten so die Chance, ggf. initiiert durch Branchenverbände, spezifische Regeln für die Sicherheit von Big Data zu entwickeln. Das Problem der fehlenden Verbindlichkeit von Verhaltensregeln könnte dadurch entschärft werden, dass Unternehmen durch die Satzung eines Branchenverbandes verpflichtet werden.[277] Auch die Zertifizierung in §§ 9, 9a BSIG könnte ein wichtiger Schritt sein, um zumindest besonders gefährdete Big-Data-Anwendungen zu bewerten. Hierzu dürfte es sinnvoll sein, Zertifizierungsschemata zu erstellen, die auf eine ganzheitliche Betrachtung einer gesamten Big-Data-Anwendung abzielen.

5.3.3 Standardisierungsmaßnahmen

Daneben könnten die Aufsichtsbehörden und entsprechende Normungsgremien bei der Erstellung von Standards, auch unter besonderer Berücksichtigung der Interoperabilität, mehr Unterstützung leisten.[278] Die im BSIG bereits angelegten branchenspezifischen Standards bilden eine gute Möglichkeit, um einerseits durch entsprechende Partizipation die Akzeptanz der verantwortlichen Akteure zu gewinnen, aber gleichzeitig auch Verpflichtungen zu präzisieren und zu verallgemeinern. Hier könnten insbesondere KMUs mit Big-Data-Anwendungen stärker dazu animiert werden, sich vermehrt an solchen Standards zu beteiligen. Des Weiteren könnte das BSI Richtlinien und Umsetzungsvorschläge für die IT-Sicherheit von Big-Data-Anwendungen, oder Teilbereiche davon, erlassen. Das BSI könnte Unternehmen dementsprechend noch stärker ermutigen, ihre Systeme abzusichern.[279]

Auch mit Blick auf die Operationalisierung von Art. 32 DSGVO wäre näher zu untersuchen, welcher Wert in einer Konkretisierung durch den Europäischen Datenschutzausschuss und/oder die nationalen Aufsichtsbehörden liegen könnte. Die Behörden haben in Deutschland und Europa inzwischen eine Vielzahl von Handreichungen erarbeitet, sodass auch der Bereich der Risiken für personenbezogene Daten in Big-Data-Anwendungen für

[277] Für die Verhaltensregeln in Art. 40 DSGVO, *Roßnagel* in *Simitis/Hornung/Spiecker gen. Döhmann*, DSGVO/BDSG, Art. 40 Rdnr. 68.
[278] *ENISA,* Big Data Security, 2015, S. 21.
[279] *ENISA,* Big Data Security, 2015, S. 26.

eine entsprechende ganzheitliche Betrachtung geeignet erscheint. Einen ähnlichen Weg beschreitet die neue NIS-2-RL, wenn sie als Vorgabe an eine nationale Cybersicherheitsstrategie vorsieht, bislang wenig oder nicht regulierte kleinere und mittlere Unternehmen zukünftig durch „Bereitstellung leicht zugänglicher Orientierungshilfen und Unterstützung für ihre spezifischen Bedürfnisse"[280] dabei zu helfen, ihr Grundniveau für Cyberresilienz und Cyberhygiene zu stärken. Der Schutz von Big-Data-Anwendungen in KMU könnte ein wichtiges Anwendungsfeld für diesen Bereich einer künftigen Cybersicherheitsstrategie sein.

Schließlich sollten Standardisierungsmaßnahmen entsprechend dem bereits in Art. 25 DSGVO verankerten Grundsatz des „Privacy by Design" sowohl die Hersteller als auch die Anwender von IT-Systemen in Big-Data-Anwendungen i.S.e. „Securiy by Design"[281] adressieren. Einen ersten Weg in diese Richtung geht der Cyber Resilience Act,[282] der Cybersicherheitsvorschriften für sichere IT-Komponenten macht. Idealerweise sollte der Grundsatz auch im BSIG verankert werden.

5.4 Zusammenfassung und Fazit

Zusammenfassend fehlt es im Datenschutz- und IT-Sicherheitsrecht an verpflichtenden und abgestimmten IT-Sicherheitsmaßnahmen, die dem besonderen Schutzbedürfnis von Big Data Rechnung tragen. Zum einen erfassen rechtliche Maßgaben aus dem IT-Sicherheitsrecht bislang und auch de lege ferenda nur einen Teil der Unternehmen, die Big-Data-Anwendungen einsetzen. Zum anderen ist Art. 32 DSGVO zwar für sämtliche Unternehmen verpflichtend umzusetzen, die personenbezogene Daten verarbeiten; jedoch werden aufgrund seiner Schutzrichtung nicht sämtliche IT-Systeme erfasst, und die fehlenden Konkretisierungen behindern eine effektive Anwendung. Erschwerend kommt hinzu, dass die o.g. heterogenen Teilbereiche der Big Data meist nur getrennt betrachtet werden. Ein holistischer Ansatz, der die Gefahren aus der Zusammenführung der technologischen Neuerungen erfasst, taucht selten auf. Daher sollten Unternehmen zum Schutz personenbezogener Daten Big-Data-Anwendungen ganzheitlich in ihrem Datenmanagementsystem und IT-Sicherheitskonzepten beurteilen.

Diese auf abstrakter Ebene liegende Zielvorstellung bedarf freilich einer erheblichen Konkretisierung, die bislang gerade im Bereich von Art. 32 DSGVO vielfach durch die Verantwortlichen selbst geleistet werden muss und v.a. im Bereich von KMU diese nicht selten vor Probleme stellt. Zwar bestehen mit Verhaltensregeln, Zertifizierungen und Standardisierungen bereits gesetzlich normierte Möglichkeiten, um den Schutz bislang kaum regulierter Unternehmen hinreichend zu gestalten. Allerdings bedarf es insoweit mehr Unterstützung für die Unternehmen, damit diese die Instrumente überhaupt nutzen und, sofern sie dazu bereit sind, effektiv umsetzen können. Neben einer umfassenderen Sensibilisierung für die Risiken von Big Data sollten daher insbesondere KMUs motiviert werden, sich an dem Schutz entsprechender IT-Systeme verstärkt zu beteiligen. Denn letztendlich können die Ziele des IT-Sicherheits- und Datenschutzrechts nur erreicht werden, wenn sämtliche Schutzziele gemeinsam berücksichtigt werden.[283]

[280] Vgl. Art. 7 Abs. 2 lit. i) NIS-2-RL.
[281] Dazu *ENISA,* Privacy by design in big data, 2015.
[282] COM/2022/454 final, https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:52022PC0454.
[283] *Deusch/Eggendorfer in Taeger/Pohl*e: Computerrecht, 50.1, Rdnr. 311.

6. Zusammenfassung

Eine umfangreiche Verarbeitung großer Mengen personenbezogener Daten (Big Data) kann mit erhöhten Risiken für die Rechte und Freiheiten betroffener Personen einhergehen. In besonderer Weise sind Verhaltensbeeinflussungen ein Problem (Kapitel 2). Gleichzeitig birgt die systematische Verarbeitung großer Datenmengen aus unterschiedlichsten Quellen und in verschiedensten Formen auch Chancen. Eine kluge Technikentwicklung und -nutzung, die im Einklang mit rechtlichen Anforderungen steht, vermag diese Risiken zu verringern, ohne die Chancen signifikant zu verändern. Zudem kann sie Rechtsunsicherheiten, die an vielen Stellen bestehen, reduzieren oder sogar vollständig vermeiden. Die Studie greift aus der Vielzahl der betroffenen Rechtsgebiete ausgewählte Fragestellungen heraus, nämlich die besondere Verantwortlichkeit des Staates (Kapitel 2), die Berücksichtigung der Rechte Dritter (Kapitel 3), datenschutzrechtliche gebotene Maßnahmen (Kapitel 4) sowie skalierbare IT-sicherheitsrechtliche Ansätze (Kapitel 5).

Die Gefahren und Risiken von Big-Data-Auswertungen und darauf basierenden Entscheidungen sind vielfältig. Verhaltensbeeinflussung ist ein häufiges Ziel. **Verhaltensbeeinflussung durch den Staat** ist nicht per se unzulässig; die Digitalisierung der Verwaltung und das grundsätzlich berechtigte Interesse des Staates, mit einer Instrumentenvielfalt Steuerung zu erzielen, verlangen, dass auch solche Instrumente geprüft und eingesetzt werden können. Sie ist aber mit besonderen Risiken verbunden, weil sich Bürger staatlicher Verhaltensbeeinflussung nur schwer entziehen können, und weil das grundsätzliche Über-/Unterordnungsverhältnis zwischen Staat und Bürger dem Staat ein rechtlich relevantes Übergewicht verleiht. Sie muss daher mindestens so ausgestaltet sein, dass sie dem Einzelnen seine Entscheidungsfreiheit belässt; ansonsten muss der Staat sich auf ordnungsrechtliche Instrumente, gegen die Rechtsschutz eindeutig möglich ist, beschränken. U.a. ist der Einsatz des seit einiger Zeit hoch gehandelten sog. „Nudging" im Hinblick auf die Entscheidungsfreiheit ambivalent zu beurteilen. Die Veränderung der Entscheidungsarchitektur soll diejenigen, die diesem Nudging ausgesetzt werden, dazu bringen, die erwünschten Entscheidungen zu treffen. Diese Veränderung wird aber nicht offen kommuniziert und ist daher auch nicht offen angreifbar. Geschehen solche staatlichen Verhaltensbeeinflussungen zusätzlich auf Grundlage der umfassenden Auswertung und Analyse personenbezogener Daten, insbesondere im Rahmen von Personalisierung, von denen der Einzelne womöglich gar nichts weiß, ist die Schlussfolgerung naheliegend, dass die Maßnahme sich zumindest auf der Schwelle von einer „bloßen" Steuerung hin zu einer (die Unwissenheit der Bürger und Nicht-Erkennbarkeit der Datenauswertung bis hin zu fehlender Beurteilungsmöglichkeit der Auswertungsprozesse nutzenden) Manipulation befindet.

Die grundlegenden Problemfelder wie insbesondere das Vorliegen des sog. „Automation Bias", nämlich einer Überbewertung der Verlässlichkeit automatisierter Ergebnisse, müssen in allen Anwendungskontexten minimiert werden. Die daraus folgenden Fragen werden sich in Zukunft immer häufiger und immer dringlicher stellen: Der Staat bedient sich nun einmal Elementen der indirekten Verhaltenssteuerung, die auf Daten basieren, und durch die Digitalisierung nehmen Datenauswertungen in allen Bereichen zu. Die Einfachheit und Verbreitung der Technik macht den Einsatz immer leichter und ubiquitärer, gleichzeitig aber auch immer schwerer erkennbar. Umso wichtiger ist es, sich jetzt mit den damit einhergehenden Problemen auseinanderzusetzen und grundrechtsverträgliche Konzepte zu entwickeln. Nur so können in einem steten Prozess zugleich die Chancen der Technologie genutzt und ihre Gefahren begrenzt werden. Zudem werden damit die Anwender erheblich von Rechtsunsicherheit und Einzelfallentscheidungen entlastet.
Zu den Aufgaben der Beseitigung von Rechtsunsicherheiten bei Big Data in großen systematischen Auswertungen zählt auch die verlässliche Absicherung der **Schutzrechte** an

betroffenen Daten. Zum einen ist zu klären, ob ohne vertraglich abgesicherte Datengewinnung eine Datenerhebung aus frei zugänglichen Internetseiten de lege lata rechtlich zulässig sein kann. Dies hängt – neben der datenschutzrechtlichen Zulässigkeit – von der möglichen Schutzfähigkeit der auszulesenden Daten sowie der Anwendbarkeit urheberrechtlicher Schrankenregelungen ab. Große Datenmengen unterschiedlichster Quellen und Formate in systematischer Auswertung finden zentralen Einsatz in Anwendungen künstlicher Intelligenz (KI), sodass auch spezielle Vorschriften für diese Systeme zu berücksichtigen sind. Haftungsrechtlich ist die Verarbeitung großer Datensätze mit der Herausforderung behaftet, dass die KI-Betreiber gegenüber Dritten de lege lata auch für mögliche Fehler in den Datensätzen haften – was übrigens für „einfache" algorithmische Systeme noch unklar ist.

Im Hinblick auf den KI-VO-E ist unabhängig von der Größe der zu verarbeitenden Datensätze der Einsatzzweck für die notwendigen Maßnahmen zur Risikobeherrschung maßgeblich. Es wird daher eine Herausforderung für KI-Betreiber sein, bei der Verarbeitung großer Datenmengen sichere vertragliche Regelungen zu finden, welche sich im Einklang mit den künftigen Regulierungen befinden, ohne den eigenen Spielraum zur Verwertung der KI-Ergebnisse zu sehr einzuengen.

Rechtsunsicherheit besteht nicht nur bei der Frage des normativen Schutzes, sondern auch dahingehend, wie Daten technisch geschützt werden können. Ein besonders prägnantes Schutzinstrument kann der Einsatz von **Anonymisierung** sein, sofern der jeweilige Kontext eine Verarbeitung rein anonymer Daten überhaupt zulässt. Jedoch birgt die Anonymisierung personenbezogener Daten für Organisationen das Risiko von Rechtsunsicherheit. Dies begründet sich nicht zuletzt durch die teils unspezifischen Anforderungen an die Anonymisierung nach der DSGVO und darüber, dass durch Re-Identifikation frühere anonymisierte Daten doch wieder einen Personenbezug aufweisen können. Die große Verfügbarkeit von Daten und die systemische Auswertung können schnell dazu führen, dass personenbezogene Datensätze einen anonymisierten Datenbestand „infizieren". Denn die eindeutige rechtliche Einordnung ist abhängig von den zunehmenden technischen Möglichkeiten der Kombinierbarkeit von Daten, die in großen Datenmengen bei systematischer Auswertung erst recht wächst. Daher bedarf es einer (technischen) Möglichkeit, den Anonymitätsgrad eines Datensatzes mit rechtlichen Konsequenzen zu bewerten.

Technische Lösungen sollen in einem solchen Fall die Einhaltung der normativen Anforderungen konkret prüfen und damit für Anwender erleichtern, die Rechtssicherheit ihres Vorgehens zu ermitteln und auch zu dokumentieren. Eine solche Möglichkeit birgt die Entwicklung eines Metrikensystems. In dieses Metrikensystem lässt sich ein zuvor anonymisierter Datensatz einfügen und testen, ob ein scheinbar anonymisierter Datensatz tatsächlich anonym ist. Dies ist insbesondere angesichts der dynamischen Entwicklung von Datensätzen von großer Bedeutung. Mithilfe von Metriken könnte beurteilt und überprüft werden, ob die Anforderungen an die Anonymität von einem anonymisierten Datensatz erfüllt werden (ggf. in Abhängigkeit vom Anwendungsfall) und ob anonymisierte Daten über die Zeit anonym bleiben.

Ein weiteres Instrument zur Vermeidung von unerwünschten Folgen der Verhaltensbeeinflussung durch den Einsatz großer Datenmengen und zur Beseitigung von Rechtsunsicherheiten ist der Einsatz von **IT-Sicherheit**. Im Datenschutz- und IT-Sicherheitsrecht fehlt es allerdings an verpflichtenden und abgestimmten IT-Sicherheitsmaßnahmen, die dem besonderen Schutzbedürfnis von Big Data Rechnung tragen, um sicherzustellen, dass Daten, Anwendungen und Systeme vertraulich, verfügbar und authentisch sind. Zum einen erfassen rechtliche Maßgaben aus dem IT-Sicherheitsrecht bislang und auch nach den aktuellen Plänen de lege ferenda nur einen Teil der Unternehmen, die Big-Data-Anwendun-

gen einsetzen. Zum anderen sind durch sämtliche Organisationen technisch-organisatorische Maßnahmen nach der DSGVO verpflichtend umzusetzen, sofern sie personenbezogene Daten verarbeiten; jedoch werden nicht sämtliche IT-Systeme erfasst. Erschwerend kommt hinzu, dass die Teilbereiche von Big Data meist nur getrennt betrachtet werden. Ein holistischer Ansatz, der die Gefahren aus der Zusammenführung der technologischen Neuerungen erfasst, taucht selten auf. Daher sollten Unternehmen zum Schutz personenbezogener Daten ihre Big-Data-Anwendungen als Einheit in ihren Datenmanagementsystemen und IT-Sicherheitskonzepten betrachten.

Diese auf abstrakter Ebene liegende Zielvorstellung bedarf freilich einer erheblichen Konkretisierung, die bislang gerade im Bereich des technisch-organisatorischen Datenschutzes nach der DSGVO vielfach durch die Verantwortlichen selbst geleistet werden muss und v.a. KMU nicht selten vor Probleme stellt. Zwar bestehen mit Verhaltensregeln, Zertifizierungen und Standardisierungen bereits gesetzlich normierte Möglichkeiten, um den Schutz bislang kaum regulierter Unternehmen hinreichend zu gestalten. Allerdings bedarf es insoweit mehr Unterstützung für die Unternehmen bei der Umsetzung der Maßnahmen. Neben einer umfassenderen Sensibilisierung für die Risiken von Big Data sollten daher insbesondere KMU motiviert werden, sich verstärkt an dem Schutz entsprechender IT-Systeme zu beteiligen. Denn letztendlich können die Ziele des IT-Sicherheits- und Datenschutzrechts nur erreicht werden, wenn sämtliche Schutzziele der beiden Rechtsgebiete und alle zu schützenden Bestandteile der IT gemeinsam berücksichtigt werden.

Die Heterogenität der Herausforderungen einschließlich der unterschiedlichen Granularität der aufgeworfenen Problembereiche erfordert eine anspruchsvolle rechtliche und technische Gestaltung. Nur diese kann sicherstellen, dass die Nutzung von Big Data in systematischer Auswertung und in systematischem Einsatz menschengerecht und fair verläuft und somit die Chancen für einen gesellschaftlichen Mehrwert verwirklicht.

Literatur

van Aaken, Anne	Constutional Limits to Paternalistic Nudging: A Proportionality Assessment. In: *Kemmerer, Alexandra/ Möllers, Christoph/ Steinbeis, Maximilian/ Wagner, Gerhard (Hrsg.)*, Choice Architecture in Democracies. Exploring the Legitimacy of Nudging, Baden-Baden 2016, S. 161-196.
Agencia Española de Protección de Datos	10 Misunderstandings related to Anonymisation, 2021, über: https://edps.europa.eu/system/files/2021-04/21-04-27_aepd-edps_anonymisation_en_5.pdf.
Auer-Reinsdorff, Astrid/Conrad, Isabell (Hrsg.)	IT- und Datenschutzrecht, 3. Auflage, München 2019.
Alexy, Robert	Theorie der Grundrechte, 7. Auflage, Frankfurt am Main 2015.
Ammann, Franz-Ernst / Sowa, Aleksandra	Systematische Entwicklung von Metriken zur Beurteilung der Datensicherheit, DuD 2012, S. 247-251.
An Coimisiún um Chosaint Sonraí	Guidance Note: Guidance on Anonymisation and Pseudonymisation, 2019, über: https://www.dataprotection.ie/sites/default/files/uploads/2020-09/190614%20Anonymisation%20and%20Pseudonymisation.pdf.
Arbuckle, Luk / El Emam, Khaled	Building an Anonymization Pipeline. Creating Safe Data, O'Reilly Media, 2020.
Article 29 Data Protection Working Party	Opinion 4/2007 on the concept of personal data (WP 136), https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf. Opinion 05/2014 on Anonymisation Techniques (WP 216), über: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.
Assion, Simon	Überwachung und Chilling Effects. In: *Telemedicus e.V. (Hrsg.)*, Überwachung und Recht, Tagungsband der Telemedicus Sommerkonferenz 2014, S. 31-82.
Baer, Franziskus	Staatliche Steuerung durch Nudging im Lichte der Grundrechte, Tübingen 2023.
Barev, Torben Jan/ Dickhaut, Ernestine/ Schomberg, Sabrina/ Janson, Andreas/ Schöbel, Sofia/ Grote, Thomas/ Hornung, Gerrit/ Leimeister, Jan Marco	Systematisches Design digitaler Privacy Nudges. Handlungsempfehlungen zur Gestaltung von digitalen Privacy Nudges, eine Handblungsbröschüre im Rahmen des BMBF-Projekts „Nudging Privacy in der digitalisierten Arbeitswelt – Systematische Konzeptentwicklung und Pilotierung (Nudger)", Kassel 2022, über: https://kobra.uni-kassel.de/bitstream/handle/123456789/14234/kup_9783737610889.pdf?sequence=1&isAllowed=y.

Barth, Susanne/ de Jong, Menno D. T.	The privacy paradox - Investigating discrepancies between expressed privacy concerns and actual online behavior –A systematic literature review, Telemedicus and Informatics 2017, S. 1038-1058.
Bischoff, Claudia	Pseudonymisierung und Anonymisierung im Rahmen klinischer Prüfungen von Arzneimitteln (Teil I), PharmR 2020, S. 309-315.
Bohannon, John	Genealogy Databases Enable Naming of Anonymous DNA Donors, Science 2013, 339 (6117), S. 262.
Borah, Abhishek/ Skiera, Bernd	Marketing and Investor Behavior: Insights, Introspections, and Indications, International Journal of Research in Marketing, Vol. 38, Issue 4, December 2021, S. 811-816.
Bomhard, David/ Siglmüller, Jonas	Europäische KI-Haftungsrichtlinie, RDi 2022, S. 506-513.
Bozdag, Engin	Bias in algorithmic filtering and personalization, Ethics and Information Technology, Ethics and Information Technology, Vol. 15, Issue 3, 2013, S. 209-227.
Breher, Nina/ Lehmann/Hendrik	Vordenker zu ChatGPT, Tagesspiegel vom 26.03.2023, über: https://www.tagesspiegel.de/chancen-und-gefahren-der-kunstlichen-intelligenz-das-ist-schon-ziemlich-revolutionar-9558341.html.
Buchmann, Johannes et al.	Digitalisierung und Demokratie, Schriftenreihe zur wissenschaftsbasierten Politikberatung, Deutsche Akademie der Naturforscher Leopoldine e.V. – Nationale Akademie der Wissenschaften, Halle (Saale) 2021.
Bundesamt für Sicherheit in der Informationstechnik.	BSI-Lagebericht 2022.
	IT-Grundschutz-Kompendium, CON.1.
Burchardi, Sophie	Risikotragung für KI-Systeme, EuZW 2022, S. 685-692.
Blum, Daniel	NSA-Affäre - Epoche der totalen Überwachung eingeläutet, über: https://www.deutschlandfunk.de/nsa-affaere-epoche-der-totalen-ueberwachung-eingelaeutet-100.html.
Büscher, Max/ Gutjahr, Amina/ Hornung, Gerrit et al.	Verhaltensbeeinflussung durch Beobachtung? Explorative Studie des Nationalen Forschungszentrums für angewandte Cybersicherheit ATHENE, 2022.
Britz, Gabriele	Einzelfallgerechtigkeit versus Generalisierung. Verfassungsrechtliche Grenzen statistischer Diskriminierung, Tübingen 2008.
Calliess, Christian/ Ruffert, Matthias (Hrsg.)	EUV/AEUV. Das Verfassungsrecht der Europäischen Union mit Europäischer Grundrechtecharta. Kommentar, 6. Auflage, München 2022.

Chiampi Ohly, Diana — SoftwareRecht: Von der Entwicklung zum Export, 4. Auflage, Frankfurt a.M. 2022.

Chibanguza, Kuuya/ Kuß, Christian/ Steege, Hans — Künstliche Intelligenz, 1. Auflage, Baden-Baden 2022.

Cormen, Thomas/ Leiserson, Charles/ Rivest, Ronald et al. — Algorithmen – Eine Einführung, 4. Auflage, München 2013.

Crawford, Kate/ Calo, Ryan — There is a blind spot in AI reseach, Nature 2016, S. 311-313.

Damberger,Thomas — Von der Abschaffung des Lehrers, Lernende Schule 20 (2017) 79, S. 22-24.

Damm, Frank/Fischer, Hans-Peter — Lieferkette: Wie Cyber-Security von adäquater Zusammenarbeit abhängt, DuD 2019, S. 418-425.

Datenethikkommission der Bundesregierung (Hrsg.) — Gutachten der Datenethikkommission der Bundesregierung, Berlin 2019, über: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethik-kommission.pdf?__blob=publicationFile&v=6.

Datenschutzkonferenz — Kurpapier Nr. 18 Risiko für die Rechte und Freiheiten natürlicher Personen, über: https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_18.pdf.

Denninger, Erhard — Die Wirksamkeit der Menschenrechte in der deutschen Verfassungsrechtsprechung, JZ 1998, S. 1129-1135.

Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit — Positionspapier zur Anonymisierung unter der DSGVO unter besonderer Berücksichtigung der TK-Branche, über: https://www.bfdi.bund.de/SharedDocs/Downloads/DE/Konsultationsverfahren/1_Anonymisierung/Positionspapier-Anonymisierung.pdf?__blob=publicationFile&v=4.

Desai, Tanvi/ Ritchie, Felix/ Welpton, Richard — Five Safes: designing data access for research. Economics Working Paper Series. University of the West of England, Bristol, England. Faculty of Business and Law, 2016.

Deutscher Bildungsserver — Künstliche Intelligenz in der Schule, 2023, über: https://www.bildungsserver.de/kuenstliche-intelligenz-in-der-schule-12990-de.html.

Deutscher Ethikrat — Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz, Stellungnahme, März 2023, über: https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf.

Die Techniker — Belohnung durch Bewegung, über: https://www.tk.de/techniker/magazin/digitale-gesund-

	heit/spezial/tk-fit-2066260?gclid=Cj0KCQjwsIejBhDOARIsANYqkD0ucJ2AOY43pHy4O8HJ5_eWT7qm6vVKQj5U2nxP1xHQoBAf1R3AnbMaAuASEALw_wcB.
Djeffal, Christian	IT-Sicherheit 3.0: Der neue IT-Grundschutz, MMR 2019, S. 289-294.
Dreier, Horst (Hrsg.)	Grundgesetz-Kommentar, Band 1, 3. Auflage, Tübingen 2013.
Dreier, Thomas/ Schulze, Gernot	Urheberrechtsgesetz, 7. Auflage, München 2022.
Dürig, Günter (Begr.)/ Herzog, Roman/ Scholz, Rupert (Hrsg.)	Grundgesetz-Kommentar, Band 1, 99. EL September 2022.
Ebers, Martin	StichwortKommentar Legal Tech, Baden-Baden 2023.
Ebers, Martin/ Steinrötter, Björn (Hrsg.)	Künstliche Intelligenz und smarte Robotik im IT-Sicherheitsrecht, Baden-Baden 2021.
Ebert, Andreas/ Spiecker gen. Döhmann, Indra	Der Kommissionsentwurf für eine KI-Verordnung der EU, NVwZ 2021, 1188-1193.
Ehmann, Eugen/ Selmayr, Martin	Datenschutz-Grundverordnung – Kommentar, 2. Auflage, München 2018.
Eidenmüller, Horst	Liberaler Paternalismus, JZ 2011, S. 814-821.
Eisele, Jörg/ Böhm, Kristine	Potential und Risiken von Predictive Policing. In: *Beck, Susanna/ Kusche, Carsten/ Valerius, Brian (Hrsg.)*, Digitalisierung, Automatisierung, KI und Recht, S. 517-534.
El Emam, Khaled/ Malin, Bradley	Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk. Washington D.C.: National Academies Press, 2015.
ENISA	Big Data Security – Good Practices and Recommendations on the Security of Big Data Systems, 2015.
Enzmann, Matthias/ Selzer, Annika/ Spychalski, Dominik	Data Erasure under the GDPR – Steps towards Compliance, EDPL 2019, S. 416-420.
Epping, Volker/ Hillgruber, Christian	Beck'scher Online-Kommentar Grundgesetz, 55. Edition, Stand 15.05.2023, München 2023.
Ernst, Christian	Algorithmische Entscheidungsfindung und personenbezogene Daten, JZ 2017, S. 1026-1036.
European Commission	Questions and Answers – Data protection reform, 21. Dezember 2015, über: https://ec.europa.eu/commission/presscorner/detail/pt/MEMO_15_6385.

European Data Protection Board	Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak, über: https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_20200420_contact_tracing_covid_with_annex_en.pdf.
	EDPB Work Programme 2021/2022, über: https://edpb.europa.eu/system/files/2021-03/edpb_workprogramme_2021-2022_en.pdf.
Ferguson, Andrew Guthrie	Big Data and Predictive Reasonable Suspicion, University of Pennsylvania Law Review 2015, S. 327-410.
Friele, Minou/ Bröckerhoff, Peter/ Fröhlich, Wiebke et al.	Digitale Daten für eine effizientere Prävention: Ethische und rechtliche Überlegungen zu Potenzialen und Risiken, Bundesgesundheitsblatt 2020, S. 741-748.
Fröhlich, Wiebke/ Spiecker gen. Döhmann, Indra	Können Algorithmen diskriminieren?, Verfassungsblog vom 26. Dezember 2018, über: https://verfassungsblog.de/koennen-algorithmen-diskriminieren/.
Gausling, Tina/ Gertz, Michael/ Martini, Mario et al.	Künstliche Intelligenz im digitalen Marketing. Datenschutzrechtliche Bewertung KI-gestützter Kommunikations-Tools und Profiling- Maßnahmen, ZD 2019, S. 335-341.
Gersdorf, Hubertus/ Paal, Boris	Informations- und Medienrecht – Kommentar, 2. Auflage, München 2022.
Gierschmann, Sybille	Gestaltungsmöglichkeiten durch systematisches und risikobasiertes Vorgehen – Was ist schon anonym? Planung und Bewertung der Risiken der Anonymisierung, ZD 2021, S. 482-486.
Gigerenzer, Gerd/ Todd, Peter M./ABC Research Group	Simple Heuristics That Make Us Smart, Oxford, 1999.
Gola, Peter/ Heckmann, Dirk (Hrsg.)	Datenschutzgrundverordnung Bundesdatenschutzgesetz Kommentar, 3. Auflage, München 2022.
Görres-Gesellschaft/ Verlag Herder (Hrsg.)	Staatslexikon – Recht, Wirtschaft, Gesellschaft, Band 3, Freiburg 2019.
Gräfe, Hans-Christian/ Kahl, Andreas	KI-Systeme zur automatischen Texterstellung, MMR 2021, S. 121-126.
Groger, Thomas/ Stock, Jürgen	Cybercrime – Herausforderung für die internationale Zusammenarbeit, ZRP 2017, S. 10-14.
Grünberger, Michael	Reformbedarf im AGG: Beweislastverteilung beim Einsatz von KI, ZRP 2021, S. 232-235.
Grunert, Frank	Paternalismus in der politischen Theorie der deutschen Aufklärung. In: *Anderheiden, Michael/ Bürkli, Peter/ Heinig, Hans Michael/ Kirste, Stephan/ Seelmann, Kurt (Hrsg.),* Paternalismus und Recht, S. 9-28.

Guijarro Santos, Victoria	Nicht besser als nichts – Ein Kommentar zum KI-Verordnungsentwurf, ZfDR 2023, S. 23-42.
Gutjahr, Amina/ Limberger, Victor	Informationelle Trennungsgrundsätze in der Sicherheitsarchitektur des 21. Jahrhunderts, DÖV 2022, S. 848-857.
Haake, Daniel	Prognose von Wohnungseinbrüchen mithilfe von Machine-Learning-Algorithmen, WISTA 2/2021, S. 59-71.
Hacker, Philipp	Immaterialgüterrechtlicher Schutz von KI-Trainingsdaten, GRUR 2020, S. 1025-1033.
Hartmann, Matthias/ Jacobsen, Jonas	„Maschinenlesbarkeit" des Rechtevorbehalts im neuen § 44b UrhG, MMR-Aktuell 2021, 441332.
Härtel, Ines	Digitalisierung im Lichte des Verfassungsrechts – Algorithmen, Predictive Policing, autonomes Fahren, LKV 2019, S. 49-60.
Heeren, Gesine	Neuronale Grundlagen der Verlustaversion, Bonn 2018.
Helmke, Jan Torben/ Link, Hendrik/ Schild, Hans-Hermann	Zertifizierungskriterien für Verarbeitungstätigkeiten, DuD 2023, S. 100-107.
Hennemann, Moritz	Datenlizenzverträge, RDi 2021, S. 61-70.
Hennemann, Moritz/ Steinrötter, Björn	Data Act – Fundament des neuen EU-Datenwirtschaftsrechts?, NJW 2022, S. 1481-1486.
Herbst, Tobias	Was sind personenbezogene Daten?, NVwZ 2016, S. 902-906.
Hetmank, Sven/ Lauber-Rönsberg, Anne	Künstliche Intelligenz – Herausforderungen für das Immaterialgüterrecht, GRUR 2018, S. 574-582.
Hilgendorf, Eric	Menschenwürde und Demütigung. Die Menschenwürdekonzeption Avishai Margalits, Baden-Baden 2013.
Hillebrand, Annette/ Niederprüm, Antonia/ Schäfer, Saskja/ Thiele, Sonja/ Henseler-Unger, Iris	Aktuelle Lage der IT-Sicherheit in KMU, 2017, über: https://www.wik.org/fileadmin/files/_migrated/news_files/WIK-Studie_Aktuelle_Lage_der_IT-Sicherheit_in_KMU_Langfassung__2_.pdf.
Hölzel, Julian	Differential Privacy and the GDPR, EDPL 2019, S. 184-196.
Hoeren, Thomas/ Sieber, Ulrich/ Holznagel, Bernd (Hrsg.)	Handbuch Multimedia-Recht, Rechtsfragen des elektronischen Geschäftsverkehrs, 58. Auflage, München, März 2022.
Hoffmann, Christian/ Luch, Anika/ Schulz, Sönke et al.	Die digitale Dimension der Grundrechte. Das Grundgesetz im digitalen Zeitalter, Baden-Baden 2015.
Hoffmann-Riem, Wolfgang	Die Governance-Perspektive in der rechtswissenschaftlichen Innovationsforschung, Baden-Baden 2011.

	Innovation und Recht – Recht und Innovation, Tübingen 2016.
	Rechtliche Rahmenbedingungen für und regulative Herausforderungen durch Big Data. In: *Ders. (Hrsg.)*, Big Data – Regulative Herausforderungen, Baden-Baden 2018, S. 9-78.
	Verhaltenssteuerung durch Algorithmen – Eine Herausforderung für das Recht, AöR 2017, S. 1-42.
Hofstetter, Yvonne	Das Ende der Demokratie, 2. Auflage, München 2016.
Honer, Mathias	Nudging: Keine Herausforderung für die Grundrechtsdogmatik, DÖV 2019, S. 940-949.
Hornung, Gerrit	Das IT-Sicherheitsgesetz 2.0: Kompetenzaufwuchs des BSI und neue Pflichten für Unternehmen, NJW 2021, S. 1985-1991.
	Erosion traditioneller Prinzipien des Datenschutzrechts durch Big Data, S. 81 – 98. In: Hoffmann-Riehm, Wolfgang: Big Data – Regulative Herausforderungen (Hrsg.), Baden-Baden 2018.
Hornung, Gerrit/ Hartl, Korbinian	Datenschutz durch Marktanreize – auch in Europa? – Stand der Diskussion zu Datenschutzzertifizierung und Datenschutzaudit, ZD 2014, S. 219-225.
Hornung, Gerrit/ Herfurth, Constantin	Datenschutz bei Big Data. Rechtliche und politische Implikationen, S. 149-184. In: König, Cristian/ Schröder, Jette/ Wiegand, Erich (Hrsg.): Big Data – Chancen, Risiken, Entwicklungstendenzen, Wiesbaden 2018.
Hornung, Gerrit/ Schallbruch, Martin (Hrsg.)	IT-Sicherheitsrecht, Baden-Baden 2021.
Hornung, Gerrit/ Wagner, Bernd	Der schleichende Personenbezug: Die Zwickmühle der Re-Identifizierbarkeit in Zeiten von Big Data und Ubiquitous Computing, CR 2019, S. 565-574.
	Anonymisierung als datenschutzrelevante Verarbeitung? Rechtliche Anforderungen und Grenzen für die Anonymisierung personenbezogener Daten, ZD 2020, S. 223-228.
Horstmann, Jan	Rechtbank Den Haag: System zur Erkennung von Sozialbetrug verstößt gegen EMRK, ZD-Aktuell 2020, 07047.
Hufen, Friedhelm	Nudging – Rechtsformen, Möglichkeiten und Grenzen der sanften Beeinflussung des Menschen durch den Staat, JuS 2020, S. 193-199.
Institute of Standards and Technology	NIST Big Data Interoperability Framework: Volume 1, Definitions.
Isensee, Josef	§ 87: Würde des Menschen. In: *Merten, Detlef/ Papier, Hans-Jürgen (Hrsg.),* Handbuch der Grundrechte in

	Deutschland und Europa, Band IV. Grundrechte in Deutschland. Einzelgrundrechte I, S. 3-136.
Johnson, Eric J./ Goldstein, Daniel	Do Defaults Save Lives?, SCIENCE 2003, Vol. 302, Issue 5649, S. 1338-1339.
Käde, Lisa	Kreative Maschinen und Urheberrecht, Baden-Baden 2021.
Kahnemann, Daniel	Schnelles Denken, langsames Denken, 22. Auflage, München 2012.
Kahnemann, Daniel/ Knetsch, Jack L./ Thaler, Richard H.	Anomalies. The Endownment Effect, Loss Aversion, and Satus Quo Bias, Journal of Economic Perspectives 1991, S. 193-206.
Kant, Immanuel	Grundlegung zur Metaphysik der Sitten. In: *Weischedel, Wilhelm (Hrsg.),* Immanuel Kant. Werkausgabe in 12 Bänden, Band VII: Kritik der praktischen Vernunft. Grundlegung zur Metaphysik der Sitten, Frankfurt am Main 2000.
Kaulartz, Markus/ Braegelmann, Tom (Hrsg.)	Rechtshandbuch Artificial Intelligence und Machine Learning, 1. Auflage, München 2020.
Kelber, Ulrich/ Leopold, Nils	Personalisierung durch Profiling, Scoring, Microtargeting und mögliche Folgen für Demokratie – Funktionsweisen und Risiken aus datenschutzrechtlicher Sicht. In: *Spiecker gen. Döhmann, Indra/ Westland, Michael/ Campos, Ricardo (Hrsg.)*, Demokratie und Öffentlichkeit im 21. Jahrhundert – Zur Macht des Digitalen, Baden-Baden 2022, S. 149-175.
Kianfar, Mina	Die Wirkung einer virtuellen Hausordnung am Beispiel des Screen Scraping, DSRITB 2014, S. 821-828.
Kind, Sonja/ Weide, Sebastian	Microtargeting: psychometrische Analyse mittels Big Data, Themenkurzpril Nr. 18 des Büros für Technikfolgenabschätzung beim Deutschen Bundestag (TAB), 2017, DOI: 10.5445/IR/1000133902.
Kipker, Dennis-Kenji	EAID: EU-Datenstrategie: Welche Auswirkungen ergeben sich für den Datenschutz?, ZD-Aktuell 2022, 04465.
Kipker, Dennis-Kinji/ Dittrich, Tilmann	Rolle der Kritischen Infrastrukturen nach dem neuen NIS-2-Umsetzungs- und Cybersicherheitsstärkungsgesetz - Nationale Besonderheiten und europäische Überformung, MMR 2023, S. 481 - 487.
Kirchhof, Gregor	Nudging – zu den rechtlichen Grenzen informalen Verwaltens, ZRP 2015, S. 136-137.
Klöhn, Lars	Kapitalmarkt, Spekulation und Behavioral Finance – Eine interdisziplinäre und vergleichende Analyse zum Fluch und Segen der Spekulation und ihrer Regulierung durch Recht und Markt, Berlin 2006.

Kirste, Stephan	Harter und weicher Rechtspaternalismus, JZ 2011, S. 805-814.
Kohpeiß, Marcel/ Schaller, Till	Systeme zur Angriffserkennung nach § 8a Abs. 1a BSIG, CR 2023, 589-595.
Kolbe, Frederike	Freiheitsschutz vor staatlicher Gesundheitssteuerung – Grundrechtliche Grenzen paternalistischen Staatshandelns, Band 70, 1. Auflage, Baden-Baden, 2017.
Konferenz der Datenschutzbeauftragten des Bundes und der Länder (DSK)	*Konferenz der Datenschutzbeauftragten des Bundes und der Länder (DSK),* Entschließung: Big Data zur Gefahrenabwehr und Strafverfolgung: Risiken und Nebenwirkungen beachten, über: https://www.datenschutzkonferenz-online.de/media/en/20150318_en_BigData.pdf.
Konferenz der Datenschutzbeauftragten des Bundes und der Länder (DSK)	Entschließung Standard-Datenschutzmodell, Version 3.0, 2022.
Kronenberger, Nadja Sylvia	Nudging als Steuerungsinstrument des Rechts, Saarbrücken 2019.
Kühling, Jürgen/ Sackmann, Florian	Irrweg Dateneigentum, ZD 2020, S. 24-30.
Kuhlmann, Simone/ Trute, Hans-Heinrich	Predictive Policing als Formen polizeilicher Wissensgenerierung, GSZ 2021, S. 103-111.
Kunz, Thomas/ Waldmann, Ulrich	ML-basierte Klassifizierung von E-Mails für die datenschutzkonforme Löschung und Archivierung, in: INFORMATIK 2022, S. 589-600, über: https://dl.gi.de/bitstreams/e3e2ff06-34b6-46c8-a7cf-bcf20779f555/download.
Laney, Doug	3D Data Management: Controlling Data Volume, Velocity, and Variety", Gartner, File Nr. 949, 2001, über: https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.
Lauscher, Anne/ Legner, Sarah	Künstliche Intelligenz und Diskriminierung, ZfDR 2022, S. 367-390.
Leupold, Andreas/ Wiebe, Andreas/ Glossner, Wiebe (Hrsg.)	IT-Recht, 4. Auflage, München 2021.
Lübbe-Wolff, Gertrude	Constitutional Limits to Health-Related Nudging – a Matter of Balancing. In: *Kemmerer, Alexandra/ Möllers, Christoph/ Steinbeis, Maximilian/ Wagner, Gerhard (Hrsg.),* Choice Architecture in Democracies, S. 247-254.

von Mangoldt, Hermann/ Klein, Friedrich/ Starck, Christian	Grundgesetz Kommentar, Band 1, 7. Auflage München 2018.
Majumdar, Mitrankur	KI als helfende Lehrkraft während des Lockdowns, eGovernment vom 23. April 2020, über: https://www.egovernment.de/ki-als-helfende-lehrkraft-waehrend-des-lockdowns-a-532bcc78c646262b120ca9bbf8771aaa/.
Marnau, Ninja	Anonymisierung, Pseudonymisierung und Transparenz für Big Data: Technische Herausforderungen und Regelungen in der Datenschutz-Grundverordnung, DuD 2016, S. 428–433.
Marnau, Ninja / Berrang, Pascal / Humbert, Mathias	Anonymisierungsverfahren für genetische Daten, DuD 2018, S. 83-88.
Martini, Mario	Algorithmen als Herausforderung für die Rechtsordnung, JZ 2017, S. 1017-1025.
Maturana, Simón	Debiasing Administration: Kognitive Heuristiken und Verzerrungen im Verwaltungsverfahren, DÖV 2022, S. 941-948.
Meier, Matthias	Verhaltenswissenschaftlich inspiriertes Verwaltungshandeln. Herausforderungen und Perspektiven zur Umsetzung staatlichen Nudgings in Deutschland, Baden-Baden 2021.
Mende, Janne	Global Governance und libertärer Paternalismus: Akteure, Normativität und Legitimität, Zeitschrift für Praktische Philosophie 2016, S. 559-598.
Ministerium für Schule und Bildung des Landes Nordrhein-Westphalen	Grundlagen, über: https://www.schulministerium.nrw/schule-bildung/schulorganisation/grundlagen.
Molavi Vasse'i, Ramak	Transparenzanforderungen an Künstliche Intelligenz, K&R 2022, Beil. 1 zu H. 7/8,S, 8-12.
de Montjoye, Yves-Alexandre / Hidalgo, César A. / Verleysen, Michel et al.	Unique in the Crowd: The privacy bounds of human mobility, Scientific Reports 2013, 3:1376, S. 1-5.
Muckel, Stefan	Wandel des Verhältnisses von Staat und Gesellschaft – Folgen für Grundrechtstheorie und Grundrechtsdogmatik, VVDStRL 2019, S. 246-286.
Mühlhoff, Rainer	Predictive Privacy: Towards an Applied Ethics of Data Analytics, Ethics and Information Technology, https://doi.org/10.1007/s10676-021-09606-x.
Müller-Quade, Jörn/ Meister, Gisela/ Holz, Thorsten/ Houdeau, Detlef/ Rieck,	Künstliche Intelligenz und IT-Sicherheit – Bestandsaufnahme und Lösungsansätze. Whitepaper aus der Plattform Lernende Systeme, 2019.

Konrad/ Rost, Peter/ Schauf, Thomas/ Schindler, Werner

von Münch, Ingo/ Kunig, Philip (Begr.) Grundgesetz-Kommentar, Band 1, 7. Auflage, München 2021.

Narayanan, Arvind/ Shmatikov, Vitaly Robust de-anonymization of large sparse datasets, IEEE Symposium on Security and Privacy 2008, S. 111-125.

Niklas, Jedrzej/ Sztandar-Sztanderska, Karolina/ Szymielewicz, Katarzyna Profiling the Unemeployed in Poland: Social and Political Implications of Algorithmic Decision Making, herausgegeben von der Fundacja Panoptykon, Warschau 2015, über: https://panoptykon.org/biblio/profiling-unemployed-poland-social-and-political-implicationsalgorithmic-decison-making.

Oermann, Markus/ Staben, Julian Mittelbare Grundrechtseingriffe durch Abschreckung? Zur grundrechtlichen Bewertung polizeilicher „Online-Streifen" und „Online-Ermittlungen in sozialen Netzwerken, Der Staat, 2013, S. 630-661.

Ohm, Paul Broken promises of privacy: Responding to the surprising Failure of Anonymization, UCLA Law Review (57) 2010, S. 1701–1777.

Orwat, Carsten Diskriminierungsrisiken durch Verwendung von Algorithmen, Eine Studie erstellt mit einer Zuwendung der Antidiskriminierungsstelle des Bundes, Berlin 2019, über: https://www.antidiskriminierungsstelle.de/SharedDocs/downloads/DE/publikationen/Expertisen/studie_diskriminierungsrisiken_durch_verwendung_von_algorithmen.pdf?__blob=publicationFile&v=3.

Paal, Boris P./ Hennemann, Moritz Big Data im Recht – Wettbewerbs- und daten(schutz)rechtliche Herausforderungen, NJW 2017, S. 1697-1701.

Paal, Boris P./ Pauly, Daniel A. (Hrsg.) Datenschutz-Grundverordnung, Bundesdatenschutzgesetz - Kommentar, 3. Auflage, München 2021.

Peters, Robert/ Bovenschulte, Marc Learning Analytics – Potenzial von KI-Systemen für Lehrende und Lernende, Themenkurzprofil Nr. 42 des Büros für Technikfolgen-Abschätzung beim Deutschen Bundestag (TAB) 2021, DOI: 10.5445/IR/1000131769.

Peuker, Enrico Verfassungswandel durch Digitalisierung. Digitale Souveränität als verfassungsrechtliches Leitbild, Thübingen 2020.

Pfeiffer, Lars/ Helmke, Jan Torben Die Digitalrechtsakte der EU (DGA, DSA, DMA, KI-VO-E und DA-E) – Teil II, ZD-Aktuell 2023, 01162.

Pfitzmann, Andreas/ Köhntopp, Marit/ Shostack Adam et. Al Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology, über: https://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.22.pdf.

Purnhagen, Kai/ Reisch, Lucia	"Nudging Germany"? Herausforderungen für eine verhaltensbasierte Regulierung in Deutschland, ZEuP 2016, S. 629-655.
Rademacher, Timo/ Perkowski, Lennart	Staatliche Überwachung, neue Technologien und die Grundrechte, JuS 2020, S. 713-720.
Raue, Benjamin	»Unberührt« – das Verhältnis von DSA zur DSM-RL und zum UrhDaG, ZUM 2023, S. 160-170.
Roos, Philipp/ Weitz, Caspar Alexander	Hochrisiko-KI-Systeme im Kommissionsentwurf für eine KI-Verordnung, MMR 2021, S. 844-851.
Roßnagel, Alexander	Datenlöschung und Anonymisierung – Verhältnis der beiden Datenschutzinstrumente nach DS-GVO, ZD 2021, S. 188-192.
	Die „Überwachungs-Gesamtrechnung" – Das BVerfG und die Vorratsdatenspeicherung, NJW 2010, S. 1238-1242.
	Pseudonymisierung personenbezogener Daten – Ein zentrales Instrument im Datenschutz nach der DS-GVO, ZD 2018, S. 243-247.
Roßnagel, Alexander/ Geminn, Christian L.	Vertrauen in Anonymisierung. Regulierung der Anonymisierung zur Förderung Künstlicher Intelligenz, ZD 2021, S. 487-490.
Roßnagel, Alexander/ Nebel, Maxi	(Verlorene) Selbstbestimmung im Datenmeer, DuD 2015, S. 455 - 459.
Roßnagel, Alexander/ Scholz, Philip	Datenschutz durch Anonymität und Pseudonymität, Rechtsfolgen der Verwendung anonymer und pseudonymer Daten, MMR 2000, S. 721-731.
Rost, Martin	Das Standard-Datenschutzmodell (SDM), 2022.
Rost, Martin	Neues vom Standard-Datenschutzmodell (SDM-V 3.0), PinG 2023, S. 94-97.
Ruschemeier, Hannah	Der additive Grundrechtseingriff, Berlin 2019.
Russell, Stuart/ Norvig, Peter	Artificial Intelligence. A Modern Approach, 4. Auflage, Boston 2019.
Sachs, Michael (Hrsg.)	Grundgesetz Kommentar, 9. Auflage, München 2021.
Sacksofsky, Ute	Autonomie und Fürsorge, KJ 2021, S. 47-61.
Sandhu, Amanpreet Kaur	Big Data with Cloud Computing: Discussions and Challenges, Big Data Mining and Analytics 2022, Vo. 5 Nr. 1, 32-40.
Schallbruch, Martin	Das IT-Sicherheitsgesetz 2.0 - neue Regeln für Unternehmen und IT-Produkte. Neue Rechtslage im IT-Sicherheitsrecht (Teil I), CR 2021, S. 450-458.

	Das IT-Sicherheitsgesetz 2.0 – Befugnisse des BSI und Schutz der Bundesverwaltung. Neue Rechtslage im IT-Sicherheitsrecht (Teil II) CR 2021, S. 516-523.
Schlehahn, Eva/ Aichroth, Patrick/ Mann, Sebastian et al.	Benefits and Pitfalls of Predictive Policing, 2015.
Schliesky, Utz/ Hoffmann, Christian/ Luch, Anika et al.	Schutzpflichten und Drittwirkung im Internet. Das Grundgesetz im digitalen Zeitalter, Baden-Baden 2014.
Schoch, Friedrich/ Schneider, Jens-Peter (Hrsg.)	Verwaltungsrecht VwVfG – Kommentar, 3. Ergl., München 2022.
Schricker, Gerhard / Loewenheim, Ulrich	Urheberrecht, 6. Auflage, München 2020.
Schöbel, Sofia Marlena/ Schomberg, Sabrina/Barev, Torben Jan et al.	Zum Datenschutz gestupst? Gestaltungsorientierte Entwicklung von Privacy Nudges vor dem Hintergrund ethischer und rechtlicher Leitlinien, in: Friedewald, Michael/ Kreutzer, Michael/ Hansen, Marit (Hrsg.), Selbstbestimmung, Privatheit und Datenschutz. Gestaltungsoptionen für einen europäischen Weg, Wiesbaden 2022, S. 369-388.
Schomberg, Sabrina/ Barev, Torben Jan/ Janson, Andreas et al.	Ansatz zur Umsetzung von Datenschutz nach der DSGVO im Arbeitsumfeld: Datenschutz durch Nudging, DuD 2019, S. 774-780.
Schuppert, Gunnar Folke	Governance und Rechtsetzung. Grundlagen einer modernen Regelungswissenschaft, Baden-Baden 2011.
	§ 16: Verwaltungsorganisation als Steuerungsfaktor. In: *Wolfgang, Hoffmann-Riem/ Schmidt-Aßmann, Eberhard/ Voßkuhle, Andreas (Hrsg.)*, Grundlagen des Verwaltungsrechts, Band I: Methoden, Maßstäbe, Aufgaben, Organisation, 2. Auflage, München 2012, S. 1067-1159.
Schwabenbauer, Thomas	Heimliche Grundrechtseingriffe. Ein Beitrag zu den Möglichkeiten und Grenzen sicherheitsbehördlicher Ausforschung, Tübingen 2013.
Seidensticker, Kai/ Bode, Felix/ Stoffel, Florian	Predictive Policing in Germany, August 2018, über: http://nbn-resolving.de/urn:nbn:de:bsz:352-2-14sbvox1ik0z06.
Selzer, Annika	Datenschutzrecht – ein Kommentar für Studium und Praxis, Stuttgart 2022.
Selzer, Annika/ Timm, Ingo	Chances and Limitations of Personal and Anonymized Data Processing, INFORMATIK 2021, S. 773-787.
	Potenziale anonymer Datenverarbeitungen nutzen, DuD 2021, S. 816-820.

Sesing, Andreas/ Tschech, Angela — AGG und KI-VO-Entwurf beim Einsatz von Künstlicher Intelligenz – Einschätzung aus der Perspektive des (Anti-)Diskriminierungsrechts, MMR 2022, S. 24-30.

Simitis, Spiros/ Hornung, Gerrit/ Spiecker gen. Döhmann, Indra (Hrsg.) — Datenschutzrecht: DSGVO mit BDSG, Nomos Großkommentar, Baden-Baden 2019.

Singelnstein, Tobias — Predictive Policing: Algorithmenbasierte Straftatprognose zur vorausschauenden Kriminalintervention, NStZ 2018, S. 1-8.

Skitka, Linda J./ Mosier, Kathleen L./ Burdick, Mark — Does automation bias decision-making? International Journal of Human-Computer Studies, (1999) 55, S. 991-1006.

Sodan, Helge — Handbuch des Krankenversicherungsrechts, 3. Auflage, München 2018.

Söbbing, Thomas — Künstliche neuronale Netze, MMR 2021, S. 111-116.

Spiecker gen. Döhmann, Indra — Das rechtliche Darstellungsgebot. Zum Umgang mit Risikoinformation am Beispiel der Datenerhebung im Bundesinfektionsschutzgesetz (IfSG). In: *Engel, Christoph/ Englerth, Markus/ Lüdemann, Jörn/ Dies. (Hrsg.),* Recht und Verhalten, Tübingen 2007, S. 133-164.

§ 20: Digitalisierung, Informationsgesellschaft, Massendaten, Künstliche Intelligenz. In: *Kischel, Uwe/ Kube, Hanno (Hrsg.),* Handbuch des Staatsrechts der Bundesrepublik Deutschland, 4. Auflage, München 2023.

§ 71: Grundrechtsfragen der Digitalisierung. In: *Kahl, Wolfgang/ Ludwigs, Markus (Hrsg.),* Handbuch des Verwaltungsrechts, Band III: Verwaltung und Verfassungsrecht, Heidelberg 2022.

Kontexte der Demokratie: Parteien, Medien, Sozialstrukturen, VVDStRL 2017, S. 9-59.

Spiecker gen. Döhmann, Indra/ Towfigh, Emanuel V. — Automatisch benachteiligt – Das Allgemeine Gleichbehandlungsgesetz und der Schutz vor Diskriminierung durch algorithmische Entscheidungssysteme, Rechtsgutachten im Auftrag der Antidiskriminierungsstelle des Bundes, April 2023, über: https://www.antidiskriminierungsstelle.de/SharedDocs/downloads/DE/publikationen/Rechtsgutachten/schutz_vor_diskriminierung_durch_KI.pdf?__blob=publicationFile&v=2.

Spindler, Gerald — Verantwortlichkeiten von IT-Herstellern, Nutzern und Intermediären, Göttingen 2020.

Staben, Julian — Der Abschreckungseffekt auf die Grundrechtsausübung. Strukturen eines verfassungsrechtlichen Arguments, Tübingen 2019.

Stiemerling, Oliver „Künstliche Intelligenz"– Automatisierung geistiger Arbeit, Big Data und das Internet der Dinge, CR 2015, S. 762-765.

Stürmer, Verena Löschen durch Anonymisieren?, ZD 2020, S. 626-631.

Stummer, Sarah Issues of Verifying Anonymity: An Overview, Proceedings der GI Informatik 2022, S. 179-194.

Sunstein, Cass R. Simpler – The Future of Government, New York 2013.

The Ethics of Nudging, Yale Journal on Regulation 2015, Vol. 32, S. 413-450.

Sunstein, Cass R./ Thaler, Richard Libertarian Paternalism is Not an Oxymoron, The University of Chicago Law Review, 2003, S. 1159-1202.

Sweeney, Latanya Simple Demographics Often Identify People Uniquely, Carnegie Mellon University, Data Privacy Working Paper 3, Pittsburgh 2000.

Sydow, Gernot/ Marsch, Nikolaus (Hrsg.) DS-GVO/BDSG - Handkommentar, 3. Auflage, München 2022.

Taeger, Jürgen/ Pohle, Jan (Hrsg.) Computerrechts-Handbuch, 37. Ergl., München 2022.

Thaler, Richard/ Sunstein, Cass R. Nudge – Wie man kluge Entscheidungen anstößt, 12. Auflage, Berlin 2017.

Ulmer-Eilfort, Constanze/ Obergfell, Eva Inés Verlagsrecht, 2. Auflage, München 2021.

Vogt, Christian/ Hennies, Patrick/Endreß, Christian/Peters, Patrick (Hrsg.) Wirtschaftsschutz in der Praxis, Wiesbaden 2022.

Völker, Jan Christoph/Schnatz, Andreas/Breyer, Jonas Chancen und Risiken von Cloud-Produkten im Unternehmen, MMR 2022, S. 427-435.

Von Faber, Eberhard/ Kohler, Arndt Die Lücke: Informationssicherheit in Systemen mit künstlicher Intelligenz: Wie Algorithmen und künstliche Intelligenz zur Gefahr für die IT-Sicherheit werden, DuD 2019, S. 434–439.

Voßkuhle, Andreas § 1: Neue Verwaltungsrechtswissenschaft. In: *Hoffmann-Riem, Wolfgang/ Schmidt-Aßmann, Eberhard/ Ders. (Hrsg.),* Grundlagen des Verwaltungsrechts, Band I: Methoden, Maßstäbe, Aufgaben, Organisation, 2. Auflage, München 2012, S. 1-63.

Voßkuhle, Andreas/ Heitzer, Sonja Grundwissen – Öffentliches Recht: Verfassungsauslegung, JuS 2023, S. 312-316.

Wagner, Gerhard/ Eidenmüller, Horst In der Falle der Algorithmen? Abschöpfen von Konsumentenrente, Ausnutzen von Verhaltensanomalien und Manipulation von Präferenzen: Die Regulierung der dunklen Seite personalisierter Transaktionen, ZfPW 2019, S. 220-246.

Wandtke, Artur-Axel/ Bullinger, Winfried Urheberrecht, 6. Auflage, München 2022.

Weber, Andreas	Digitalisierung – Machen! Machen! Machen!: Wie Sie Ihre Wertschöpfung steigern und Ihr Unternehmen retten, Wiesbaden 2017.
Weichert, Thilo	Big Data und Datenschutz. Chancen und Risiken einer neuen Form der Datenanalyse, ZD 2013, S. 251-259.
Wilmer, Thomas	Rechtsfragen bei Dall-E & Co – Schutzfähigkeit der „Promptografie"?, K&R 2023, S. 385-395.
Winter, Christian/ Battis, Verena/ Halvani, Oren	Herausforderungen für die Anonymisierung von Daten Technische Defizite, konzeptuelle Lücken und rechtliche Fragen bei der Anonymisierung von Daten, ZD 2019, S. 489-493.
Wischmeyer, Thomas	Regulierung intelligenter Systeme, AöR 2018, S. 1-66.
Wissenschaftlicher Dienst des Deutschen Bundestags	Aktueller Begriff – Big Data, über: https://www.bundestag.de/resource/blob/194790/c44371b1c740987a7f6fa74c06f518c8/big_data-data.pdf.
Wolff, Johanna	Eine Annäherung an das Nudge-Konzept nach Richard Thaler und Cass R. Sunstein aus rechtswissenschaftlicher Sicht, RW 2015, S. 194-222.
Wolff, Amadeus/ Brink, Stefan/ v. Ungern-Sternberg, Antje (Hrsg.)	BeckOK Datenschutzrecht – Kommentar, 44. Edition, München 2023.
Zech, Herbert	Künstliche Intelligenz und Haftungsfragen, ZfPW 2019, S. 198-219.

ATHENE
Nationales Forschungszentrum
für angewandte Cybersicherheit