
Friedewald | Roßnagel | Heesen | Krämer | Lamla [Hrsg.]

# Künstliche Intelligenz, Demokratie und Privatheit

Nomos

**Privatheit und Selbstbestimmung in der digitalen Welt**

**Privacy and Self-Determination in the Digital World**

herausgegeben von | edited by
Dr. Michael Friedewald
Prof. Dr. Alexander Roßnagel

Band | Volume 1

Michael Friedewald | Alexander Roßnagel
Jessica Heesen | Nicole Krämer | Jörn Lamla [Hrsg.]

# Künstliche Intelligenz, Demokratie und Privatheit

Nomos

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Gestaltung Titelmotiv: Magdalena Vollmer

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

1. Auflage 2022



Publiziert von
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestraße 3–5 | 76530 Baden-Baden
www.nomos.de

Gesamtherstellung:
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestraße 3–5 | 76530 Baden-Baden

ISBN (Print): 978-3-8487-7327-5
ISBN (ePDF): 978-3-7489-1334-4

DOI: https://doi.org/10.5771/9783748913344

Onlineversion
Nomos eLibrary

# Vorwort

Um im interdisziplinären Dialog die Auswirkungen von Datafizierung, Überwachung und Künstlicher Intelligenz auszuloten und zu diskutieren, veranstaltete das vom Bundesministerium für Bildung und Forschung (BMBF) geförderte „Forum Privatheit und selbstbestimmtes Leben in der digitalen Welt“ (http://www.forum-privatheit.de) am 18. und 19. November 2021 in Wiesbaden die Konferenz „Auswirkungen der Künstlichen Intelligenz auf Demokratie und Privatheit“. Der vorliegende Band präsentiert die wichtigsten Vorträge und reflektiert die dort angestoßenen Diskussionen.

Das „Forum Privatheit“ arbeitet seit nunmehr acht Jahren – ausgehend von technischen, juristischen, ökonomischen sowie geistes- und gesellschaftswissenschaftlichen Ansätzen – an einem interdisziplinär fundierten, zeitgemäßen Verständnis von Privatheit und Selbstbestimmung. Hieran anknüpfend werden Konzepte zur (Neu-)Bestimmung und Gewährleistung informationeller Selbstbestimmung und des Privaten in der digitalen Welt erstellt. Es versteht sich über seine Kerndisziplinen hinaus als eine Plattform für den fachlichen Austausch und erarbeitet Orientierungswissen für den öffentlichen Diskurs in Form wissenschaftlicher Publikationen, Tagungen, White- und Policy-Paper.

Seit 2021 ist das “Forum Privatheit” das zentrale Begleitprojekt der vom BMBF initiierten Plattform Privatheit und wird vom Fraunhofer-Institut für System- und Innovationsforschung (ISI) in Karlsruhe und der Projektgruppe verfassungsverträgliche Technikgestaltung (provet) an der Universität Kassel koordiniert. In der Plattform Privatheit werden die vom BMBF geförderten Projekte zu den Themen Privatheit, Datenschutz und Selbstbestimmung zusammengefasst. Ziel des “Forum Privatheit” ist es, allen Bürgerinnen und Bürgern einen reflektierten und selbstbestimmten Umgang mit ihren Daten, technischen Geräten und digitalen Anwendungen zu ermöglichen. Das “Forum Privatheit” bereitet aktuelle Forschungsergebnisse für Zivilgesellschaft, Politik, Wissenschaft und Wirtschaft auf und berät deren Akteure zu ethischen, rechtlichen und sozialen Aspekten von Privatheit, Datenschutz und informationeller Selbstbestimmung.

Die Organisation der Konferenz erfolgte zusammen mit dem Hessischen Beauftragten für Datenschutz und Informationsfreiheit (HBDI). Die inhaltliche Gestaltung erfolgte zusammen mit dem ebenfalls durch das BMBF geförderten Projekt “PRIvatheit, Demokratie und Selbstbestim-

mung im Zeitalter von KI und Globalisierung" (PRIDS), an dem neben dem Fraunhofer ISI und der Universität Kassel u.a. auch noch die Universität Duisburg-Essen und das Internationale Zentrum für Ethik in den Wissenschaften der Universität Tübingen beteiligt sind.

Als Herausgeber:innen freuen wir uns, nun diesen Konferenzband präsentieren zu können. Wir danken insbesondere den Autor:innen für die Überarbeitung ihrer Vorträge und die Beisteuerung der jeweiligen Fachaufsätze. Ebenso zum Dank verpflichtet sind wir allen Beteiligten am „Forum Privatheit" sowie den Kolleg:innen, die die in diesem Band veröffentlichten Texte begutachtet haben. Die Konferenz "Auswirkungen der Künstlichen Intelligenz auf Demokratie und Privatheit" wäre ohne die vielfältige Unterstützung durch das interdisziplinäre Kollegium nicht möglich gewesen. Wir danken insbesondere all jenen, die organisatorisch oder inhaltlich an der Vorbereitung und Durchführung der Konferenz mitgewirkt haben, darunter vor allem Susanne Ruhm, Greta Runge, Frank Ebbers, Murat Karaboga und Marleen Georgesohn (Fraunhofer ISI) sowie Christian Geminn, Tamer Bile und Carsten Ochs (Universität Kassel). Darüber hinaus danken wir Barbara Ferrarese (Fraunhofer ISI) für die professionelle Wissenschaftskommunikation, Miriam Janke (Fusionistas) für die konzeptionelle Beratung und lebendige Moderation sowie Magdalena Vollmer für die kreative Live-Visualisierung der Vorträge. Prof. Dr. Ina Schieferdecker (BMBF) danken wir für die gelungene Konferenz-Eröffnung und thematische Einordnung.

Dem hessischen Landtag verdanken wir, dass wir unsere Veranstaltung in den prächtigen Räumlichkeiten des Wiesbadener Stadtschlosses durchführen durften. Der Vize-Präsidentin des Landtags Karin Müller danken wir für die herzliche Begrüßung und historische Einführung. Der Pressestelle des HBDI, insbesondere Maria Christina Rost, danken wir für die tatkräftige Unterstützung sowie produktive Zusammenarbeit.

Dieser aus der Konferenz hervorgegangene Band wäre nicht ohne tatkräftige Unterstützung bei der Manuskriptbearbeitung und -korrektur zustande gekommen. Wir möchten uns sehr herzlich bedanken bei den Kollegen, die die Begutachtung der Tagungsbeiträge übernommen haben. Für die angenehme und zielführende Zusammenarbeit mit dem Nomos-Verlag danken wir Dr. Sandra Frey.

Last but not least möchten wir uns besonders bei Dr. Heike Prasse und Kai Enzweiler (BMBF) für die Förderung des Projektverbunds sowie die engagierte Unterstützung unserer Forschungsthemen bedanken. Auch danken wir ausdrücklich Jan-Ole Malchow, der für den Projektträger VDI/VDE-IT die Forschungsarbeiten des „Forum Privatheit", die Vorbereitung der Konferenz und das Erscheinen des Bandes konstruktiv begleitet hat.

*Die Herausgeber:innen*
*Karlsruhe, Kassel, Tübingen, Duisburg, im Juli 2022*

# Inhalt

# Geleitwort

*Ina Schieferdecker, Bundesministerium für Bildung und Forschung*

Zur Jahreskonferenz 2021 hatte das Forum Privatheit in den Hessischen Landtag eingeladen – einem symbolträchtigen Ort, an dem in Hessen 1970 das weltweit erste Datenschutzgesetz verabschiedet worden ist.

Seither wird um notwendige als auch hinreichende Ausprägungen von Datenschutz gerungen, in den letzten Jahren vermehrt um den Datenschutz im digitalen Raum. Digitale Technologien durchdringen unser Leben. Dabei ist Digitalisierung neben den riesigen Potenzialen für Fortschritt, Wohlstand und Innovation zu einer beständigen Herausforderung für die Weiterentwicklung unserer Gesellschaft geworden.

Digitalisierung soll wie jede andere Technik das Leben und Arbeiten von uns Menschen erleichtern. Sie muss dazu an unseren gesellschaftlichen Bedarfen orientiert und an uns ausgerichtet sein. In der heutigen Wirtschaft dominieren jedoch häufig immer noch Fragen der Gewinnmaximierung, wobei Sekundär- und Tertiäraspekte sozialer und ökologischer Nachhaltigkeit nicht eingepreist sind. Hier kann und muss steuernd eingegriffen werden.

So muss auch durch digitale Technik die Würde des Menschen als wesentliche Zielbestimmung und als zentrales Fundament jedweden Handelns gewahrt bleiben. So müssen digitale Systeme, Anwendungen und Dienstleistungen durch den Menschen beherrschbar und handhabbar sein. In der digitalen Transformation geht es deshalb auch darum, die Wahrung der Rechte im Umgang mit digitalen Medien oder sozialen Plattformen sicherzustellen. Hierzu gehört ebenso die Durchsetzung bestehender Rechte zur Privatsphäre, Meinungsfreiheit oder zum Datenschutz im Cyberraum – auch in neuen digitalisierten Räumen, die durch das Internet der Dinge und smarte Geräte eröffnet werden.

Genau bei diesen Werten setzt der europäische Weg an. Anders als andere Regionen der Welt verbinden wir in Deutschland und Europa den digitalen Fortschritt mit Datenschutz, Meinungsfreiheit und dem Recht auf Privatheit unter Achtung der Menschenwürde und der Grundrechte. Dieser europäische Weg wird aber nur dann auch in Zukunft möglich sein, wenn wir technologisch souverän bleiben und unsere Fähigkeit zur kooperativen Gestaltung und Mitgestaltung von Schlüsseltechno-

logien und technologiebasierten Innovationen ausbauen. Technologische Souveränität umfasst die Formulierung von Anforderungen an Technologien, Produkte und Dienstleistungen entsprechend der eigenen Werte und deren Absicherung auch durch die Mitbestimmung entsprechender Standards in globalen Märkten. Vertrauen in digitale Lösungen entsteht da, wo die eingesetzte Soft- und Hardware verstanden wird und wo die Einhaltung der Anforderungen, etwa zur IT-Sicherheit, überprüfbar sind und überprüft werden. Gleichwohl geht es dabei um Produkte und Dienstleistungen, die gebraucht, die gekauft und genutzt werden, die sich im Markt durchsetzen können – und so Arbeit, Arbeitsplätze und Wohlstand sichern helfen. Was nützt die beste vertrauenswürdige Lösung, wenn diese nicht angenommen wird und letztlich weniger vertrauenswürdige Lösungen zur Anwendung kommen? Hier muss im engen Schulterschluss von Wirtschaft, Wissenschaft und Gesellschaft klug agiert und im Interesse aller ein breites Verständnis von Vertrauenswürdigkeit und anderen Qualitäten technischer Lösungen erzeugt werden. Das ist kein Selbstläufer, sondern erfordert ein gut auszubalancierendes Vorgehen.

Deutschland und Europa gehen diese anstehenden Aufgaben bereits verstärkt an und können zum Vorbild für eine digitale Gesellschaft werden. Wir Europäerinnen und Europäer sind dabei nicht nur Nutzende digitaler Technologien, sondern ebenso deren Gestalter und Entwickler.

Die Politik stellt dazu immer wieder wichtige Weichen, wie zum Beispiel mit der DSGVO oder der europäischen Datenstrategie. Aktuell wird am AI Act gearbeitet, der Europa zum Zentrum für innovationsstarke, vertrauenswürdige, KI-basierte Systeme machen soll.

Hier kommt der Wissenschaft eine große Verantwortung zu. Nicht alles, was wissenschaftlich oder technisch umsetzbar ist, ist auch sinnvoll oder erstrebenswert. Und so ist es die Aufgabe der Wissenschaft, das Verständnis der digitalen Transformation zu vertiefen und dieses Verständnis in die Breite zu tragen – und dabei auf die Chancen als auch die Risiken des digitalen Wandels hinzuweisen. Der fortschreitende wissenschaftliche Erkenntnisgewinn hilft uns, erstrebenswerte Entwicklungen zu befördern und Fehlentwicklungen zu vermeiden.

Dieser Aufgabe haben sich die Mitglieder des Forums Privatheit in besonderer Weise verpflichtet. Mit einem disziplinenübergreifenden Ansatz, der sozialwissenschaftliche, psychologische, rechtliche, ökonomische und nicht zuletzt technische Perspektiven vereint, wurde in den vergangenen sieben Jahren eine neue, ganzheitliche Herangehensweise verfolgt.

Das Forum Privatheit wirkt weit über die Grenzen des Forschungsverbundes hinaus: Es liefert regelmäßig wichtige Impulse für den gesellschaftlichen Diskurs und die weitere Technikentwicklung, beispielsweise in

Themen wie Datenschutz in der Blockchain, Privatheit und Kinderrechte, Tracking von Nutzerinnen und Nutzern im Netz – und auch beim Thema KI.

Auf der Jahreskonferenz 2021 wurde das Thema „Auswirkungen der Künstlichen Intelligenz auf Demokratie und Privatheit" in den Fokus gerückt. KI hat sich zu einer Schlüsseltechnologie unserer Zeit entwickelt. Sie bietet ganz neue Chancen und Möglichkeiten. Durch moderne Verfahren des maschinellen Lernens stehen uns bei der Auswertung umfangreicher Daten neue Qualitäten und Quantitäten beim Erkennen, Einordnen und Schlussfolgern zur Verfügung. Dies eröffnet neue Lösungsmöglichkeiten und Innovationen in Anwendungskontexten wie der Gesundheit, Mobilität oder der Sicherheit.

Diese Potentiale gehen mit Herausforderungen einher: So kann KI zur Verstärkung von Ungleichbehandlungen führen. Sie kann wie jede andere Technik missbräuchlich genutzt werden. So können Grenzen zwischen Äußerungen von Menschen und Social Bots verschwimmen, da sich Social Bots mittels KI dem Verhalten echter Nutzer annähern. Und so gibt es beispielsweise intensive Diskussionen darum, ob von Algorithmen generierte Inhalte auch als solche kenntlich gemacht werden sollten. Aber was genau ist ein algorithmengenerierter Inhalt? Wo beginnt er, wo hört er auf? Und wie können solche Inhalte kenntlich gemacht und das Kenntlichmachen wiederum abgesichert werden?

Und so ist und bleibt es wichtig, die weitere Entwicklung proaktiv mitzugestalten. Ihnen als Forschenden des Forums Privatheit kommt dabei die Aufgabe zu, aus Ihren Erkenntnissen die richtigen Impulse zu entwickeln, die dabei helfen, die Entscheidungshoheit der Menschen in den Mittelpunkt zu rücken und neue Entwicklungen zielgerichtet an den gesellschaftlichen Bedarfen auszurichten. Sie müssen die Auswirkungen von KI auf Privatheit und Demokratie unbedingt weiter in Breite und Tiefe diskutieren. Die Aufgaben werden nicht kleiner werden: Die digitale Transformation wird Jahrzehnte benötigen. Es werden immer wieder neue Fragestellungen auftreten. Dabei muss es gelingen, und ist es gerade demokratischen Gesellschaften immer wieder gelungen, Technologien einzuhegen, um Fehlentwicklungen zu begrenzen oder zu vermeiden.

Damit das gelingt, brauchen wir ebenso eine zielführende Forschungspolitik. Dem Bundesministerium für Bildung und Forschung ist das Forum Privatheit ein wichtiges Anliegen. Wir haben die Förderung deshalb nicht nur fortgeführt, sondern bereiten aktuell den Ausbau des Forums Privatheit zur „Plattform Privatheit" vor. Zukünftig wollen wir die wissenschaftliche Auseinandersetzung mit dem Thema Privatheit unter einer entsprechenden Rahmenbekanntmachung fördern. Die dynamischen Ent-

wicklungen relevanter Forschungsthemen können so schneller und flexibler adressiert werden. Ein erstes Projekt zum Thema „Privatheit, Demokratie und Selbstbestimmung im Zeitalter von Künstlicher Intelligenz und Globalisierung“ ist bereits gestartet.

Zudem hat die Bundesregierung das Thema Privatheit in ihrer Cybersicherheitsstrategie verankert. Und ebenso zentral ist es für das Forschungsrahmenprogramm „Digital. Sicher. Souverän.“ Mit diesem Programm setzen wir den Rahmen für eine Forschung, die den europäischen Weg in der Digitalisierung vorantreibt und die technologische Souveränität stärkt. Unser Ziel ist es, mit einer Forschung europäischer Prägung Innovationen anzustoßen und die technologische Souveränität Deutschlands und Europas in Zukunft zu wahren und in wichtigen Schlüsselbereichen auszubauen. Deshalb stellen wir für die Umsetzung des Programms bis 2026 mindestens 350 Millionen Euro bereit.

Mit dem „Forschungsnetzwerk Anonymisierung für eine sichere Datennutzung“ werden künftig zudem Fragen der Anonymisierung und des technischen Datenschutzes gebündelt. Der Kern dieses Netzwerks wird aus Kompetenzclustern zu wichtigen Anwendungsbereichen für die Anonymisierung von personenbezogenen Daten wie Medizin oder Mobilität bestehen. Und das Forum Privatheit sowie die zukünftige Plattform Privatheit werden auch weiterhin als wichtige Stimmen den öffentlichen Diskurs zu den Themen Privatheit und Datenschutz anregen.

Die Wahrung von Datenschutz und Privatheit nach europäischen Standards bei der Gestaltung und Entwicklung neuer und nachhaltiger Technologien ist kein Hemmschuh. Richtig aufgesetzt sind sie Quellen der Innovation. Und sie haben das Potenzial, Wirtschaft und Gesellschaft nachhaltig entsprechend unserer Werte weiterzuentwickeln.

Privatheit ist und bleibt ein zentraler Wert in unserem Wertekanon und in unseren Demokratien. Diesen Wert gilt es auch in einer digitalisierten Welt zu erhalten, zu pflegen und zu schützen. Hierfür ist interdisziplinäre Forschung ein zentraler Schlüssel. Denn: Technikentwicklung und deren kritische Begleitung müssen Hand in Hand gehen.

Für diese wichtige kritische Begleitung der digitalen Transformation danke ich allen am Forum Privatheit Beteiligten und wünsche Ihnen für den weiteren Weg und Ihre weitere Arbeit gutes Gelingen.

# Einleitung: Künstliche Intelligenz, Demokratie und Privatheit

*Michael Friedewald und Alexander Roßnagel*

## *Zum Thema dieses Bandes*

Die digitale Transformation von Gesellschaften weltweit hat in den letzten Jahren nicht nur weiter an Dynamik gewonnen, sondern auch immer deutlicher spürbar globale Wirkungs- und Problemzusammenhänge ausgebildet. Heute sind es vor allem allgegenwärtige Systeme der Künstlichen Intelligenz (KI), die im Zentrum des wissenschaftlichen, politischen, ökonomischen, normativen und regulatorischen Interesses stehen. Von besonderer Bedeutung sind hier algorithmische Datenauswertungen zur Steuerung wirtschaftlichen und gesellschaftlichen Verhaltens, die eine Bedeutung für die politische Entscheidungsfindung und die Strukturierung öffentlicher Kommunikation haben und so die Lebenswirklichkeit der Bürgerinnen und Bürger mitgestalten.

Die heute diskutierten KI-Systemen sind überwiegend Vertreter der so genannten „schwachen KI", bei der es darum geht, einzelne kognitive Fähigkeiten, vor allem Erkennen und Klassifizieren innerhalb eines engen Aufgabenbereichs in einem Computersystem nachzubilden. Eine solche Nachbildung bestimmter, als „intelligent" bezeichneter Funktionen umfasst aber kein Verständnis für die dahinterliegenden Konzepte. Die dazu heute meist genutzten Verfahren sind statistischer bzw. probabilistischer Natur, die auf einer Modellierung des betrachteten Problems basieren und weitgehend nicht durch einfache Regeln erklärt werden können. Zur Erstellung der Modelle und das „Training" der Funktionalität werden in der Regel große Datenbestände benötigt, so dass die Voraussagen, Klassifizierungen oder Entscheidungen einer KI höchstens so gut sein können wie die Qualität der „Trainingsdaten". Solche, auf „maschinellem Lernen" basierende Anwendungen haben in den letzten Jahren erheblich an (technischer) Reife gewonnen.

Unternehmen und Politik betrachten KI seit einigen Jahren als so genannte Schlüsseltechnologie und hegen hohe Erwartungen an die Möglichkeiten der ökonomischen Verwertung und administrativen Nutzung

zu Zwecken des Gemeinwohls.[1] Andere warnen eher vor den disruptiven ökonomischen Effekten und den unintendierten Folgen dieser gar nicht mehr so neuen Technologie für Gesellschaft und Demokratie. Auf der nationalstaatlichen Regulierungsebene ist es nach wie vor schwierig, die damit einhergehenden Herausforderungen in den Griff zu bekommen. Unter dem Eindruck einer „überwachungskapitalistischen" Implementierung von KI-Systemen einerseits und „überwachungsstaatlichen" Verwendung solcher Systeme andererseits stehen Selbstbestimmung und Privatheit als Grundwerte der demokratischen Gesellschaft einmal mehr vor einer Bewährungsprobe. Auch die Meinung in der deutschen Bevölkerung bildet diese beiden Pole ab. Laut einer Umfrage des Branchenverbands BITKOM aus dem Jahr 2021 betrachten über 70 % der deutschen Bürgerinnen und Bürger KI vor allem als Chance, während immerhin fast 30 % die Risiken überwiegen sieht.[2]

Die mit der KI entstehenden Formen der Datafizierung ändern nicht nur die zum Schutz von Privatheit und Selbstbestimmung erforderlichen Konzepte, sondern stellen auch das Verständnis und den Stellenwert von Privatheit und Selbstbestimmung selbst in Frage. Bislang wurde ihr Wert meist so begründet, dass Privatheit und Selbstbestimmung den Einzelnen vor illegitimer Beobachtung, Einflussnahme und Fremdbestimmung schützen und dadurch eine Grundlage für individuelle Autonomie, Selbstverwirklichung sowie freie Meinungs- und Willensbildung bieten soll.

Negative Einflüsse wurden entsprechend an überwachend oder „manipulativ" wirkenden Technologien festgemacht. Verwiesen sei an dieser Stelle auf Schlagworte wie „Gesichtserkennung", „intelligente Videoüberwachung", „Big Nudging", „Micro Targeting", „Predictive Policing" und ähnliche Nutzungsformen der KI. Tatsächlich bringen derartige Technologien und die damit einhergehenden Datenverarbeitungen in zunehmendem Maße neue, auch gruppenbezogene und gesamtgesellschaftliche Risiken mit sich. Während beispielsweise die von einer personenbezogenen Datenverarbeitung konkret Betroffenen immerhin verschiedene rechtliche Möglichkeiten zur Durchsetzung ihrer Rechte offenstehen, können sich die Mitglieder einer algorithmisch generierten Gruppe weder über ihre Zughörigkeit zu dieser Gruppe noch über die sie persönlich betreffenden Auswirkungen im Klaren sein. Möglich wird eine solche Zuordnung,

1 Vgl. z.B. die KI-Strategie der Bundesregierung. https://www.ki-strategie-deutschland.de/home.html.

2 https://www.bitkom.org/Presse/Presseinformation/Kuenstliche-Intelligenz-als-Chance (zuletzt zugegriffen: 06.07.2022)

wenn Datenverarbeitungen zunächst auf konkret zu einer natürlichen Person zuordenbare Daten verzichten und stattdessen nicht-personenbezogene Daten (bestimmte Nutzungs- oder Verhaltensweisen bzw. Attribute) als Bezugspunkt nehmen. Durch einen solche Verarbeitung der Daten werden etwa aus Surfgewohnheiten einzelner Individuen Informationen gewonnen, die in der Folge dann zur Personalisierung von Werbung oder Newsfeeds eingesetzt werden können. Indem derartige Verfahren oft jenseits etablierter Schutzkonzepte operieren, weil statistische Verfahren häufig nicht mit „personenbezogener Daten" im datenschutzrechtlichen Sinne arbeiten, laufen die Regelungen des Datenschutzes ins Leere. Künstliche Intelligenz ermöglicht so nicht nur algorithmengestützte Entscheidungen, die zur Steuerung und Organisation sozialer Systeme verwendet werden, sondern auch die Extraktion „emergenter", privater Informationen aus „unverdächtigen" Datensätzen.

Ein anderes Beispiel möglicher gesellschaftlicher Auswirkungen der KI: Wird KI auch zur Entwicklung von Social Bots genutzt, damit diese computergenerierten virtuellen Gesprächspartner möglichst menschenähnlich auftreten, kann dies die Auseinandersetzung über politische Meinungen oder soziale Haltungen wesentlich verändern. Während der Einsatz von Social Bots im Falle der Beantwortung einfacher Kundenfragen noch sinnvoll erscheint, ermöglicht dieselbe Technologie, den Diskussionsteilnehmer in politischen Auseinandersetzungen vorzugaukeln, dass reale Menschen eine bestimmte Meinung vertreten. Indem Bots in Posts oder ähnlichen Äußerungen Zustimmung oder Ablehnung zu einem Vorschlag oder einer Haltung zum Ausdruck bringen, können sie im demokratischen Diskurs Mehrheiten verändern oder bestimmten Meinungen „zum Durchbruch verhelfen". Auf diese Weise kann mit ihrer Hilfe der Effekt ausgenutzt werden, dass viele Menschen Teil der Mehrheit sein wollen und daher der von Bots vertretenen Meinung zustimmen. Mittels des Einsatzes von „Bot-Armeen" sind auf diese Weise sogar großflächige Meinungsmanipulationen möglich.

In diesem Zusammenhang ist auch die für Gesellschaft und Individuen ausgehende und zunehmende Gefahr von Deepfakes und vergleichbaren manipulativen Verfahren einzuordnen. Mittels spezieller künstlicher neuronaler Netzwerke (so genannte „generative adervsarial networks") ist es heute bereits möglich, authentisch wirkende Fälschungen von (Bewegt-)Bild- und Audiomaterial zu generieren. Mittels der auf diese Weise generierten Deepfakes können sich für Individuen Konsequenzen für ihre Privatsphäre entfalten, die sich derzeit insbesondere in Form von Rachepornographie äußern. Die möglichen Verletzungen gesellschaftlicher Werte reichen allerdings weit über das Individuum hinaus, wenn sie bei-

spielsweise zur Manipulation und Irritation politischer Prozesse verwendet werden – wie etwa die gefälschten Anrufe des Kiewer Bürgermeisters Vitali Klitschko bei europäischen Politikern im Juni 2022 gezeigt haben.

Alle diese Technologien können zu einer Gefahr für demokratische Werte werden, wenn etwa Filterblasen zur übermäßigen Verbreitung von Miss- oder Desinformation sowie zu Radikalisierungstendenzen im öffentlichen Diskurs beitragen. Illegitime Informationsbestände, die jedoch eine besonders hohe Popularität unter den Nutzenden sozialer Netzwerke genießen, entfalten häufig eine stärkere Wirkung als Richtigstellungen oder differenzierte und ausgewogene Informationsbestände. Indem Algorithmen die Aussendung von Inhalten steuern, können sie derartige soziale Verhaltensweisen bestärken und zu einer Verschärfung des Problems führen.

Solche Praktiken adressieren in der Regel alle Bevölkerungsgruppen. Es muss aber berücksichtigt werden, dass die Folgen für die Selbstbestimmung aufgrund unterschiedlicher individueller Voraussetzungen für unterschiedliche gesellschaftliche Gruppen verschieden sein können. So ist davon auszugehen, dass es sich etwa bei Kindern und Jugendlichen oder bei älteren Personen um Gruppen handelt, die gegenüber ausforschenden und verhaltenssteuernden Technologien besonders verletzlich sind, da sie auf anderen Kompetenzniveaus agieren, als Gruppen mit höherer „digital literacy". Die Fähigkeiten, Kenntnisse oder Mittel, die diesen Gruppen zum wirksamen Schutz ihrer informationellen Selbstbestimmung zu Verfügung stehen, müssen daher anders bewertet, gefördert und kollektiv abgestützt werden als im Falle der übrigen Gesellschaftsmitglieder. Darüber hinaus ist auch zu berücksichtigen, dass sich Menschen und ihr Umfeld über ihre Lebensspanne erheblich ändern und damit auch die Aussagekraft der über sie gesammelten Daten.

Die aus der Tagung des „Forum Privatheit" im November 2021 hervorgegangenen und in diesem Band gesammelten Beiträge drehen sich entsprechend um die Frage, welche Auswirkungen „Künstliche Intelligenz" auf Privatheit, auf das Recht auf informationelle Selbstbestimmung und auf demokratische Strukturen und Prozesse haben kann und wie diese zu bewerten sind. Darauf aufbauend wird thematisiert, mit welchen Mitteln – von der Regulierung über ökonomische Anreize und soziale Praktiken bis zur Technikgestaltung – auf diese Herausforderungen reagiert werden kann, um eine zukunftsgerechte Gewährleistung von Selbstbestimmung und demokratischer Teilhabe zu gewährleisten.

## *Die Beiträge*

Dieser Band gliedert sich in fünf Teile, die verschiedene Aspekte des Themenspektrums aus unterschiedlicher Perspektive und mit unterschiedlicher Schwerpunktsetzung aufgreifen.

### *Künstliche Intelligenz und Selbstbestimmung*

Die Beiträge in Teil I gehen der Frage nach, in welcher Weise KI – sowohl vom theoretischen Konzept als auch von der Umsetzung her – einen Paradigmenwechsel in der Informationsverarbeitung bewirkt. Dabei steht im Vordergrund, welche neuen Herausforderungen sich damit für individuelle und gesellschaftliche Werte, insbesondere die Selbstbestimmung stellen.

*Rainer Mühlhoff* (Universität Osnabrück) argumentiert in seinem Kapitel, dass die zentrale Herausforderung des Datenschutzes im Zeitalter von KI darin liegt, die Vorhersage sensibler Informationen über Menschen und Gruppen rechtlich zu adressieren. Denn die „prädiktive Analytik" mache es möglich, aus der Verknüpfung von Verhaltensdaten (z. B. Nutzungs-, Tracking- oder Aktivitätsdaten) mit (überwiegend) anonymen oder anonymisierten Daten viele weitere Aussagen über persönliche Eigenschaften differenzierter Gruppen von Menschen zu machen – etwa über Kaufkraft, Geschlecht, Alter, sexuelle Orientierung, ethnische Zugehörigkeit etc. Dadurch hätten die Daten anderer Menschen Auswirkungen auf einen selbst und die eigenen Daten Auswirkungen auf andere Menschen – auch wenn die Daten als „nicht personenbezogene" Daten verarbeitet werden. Indem die nachfolgende gesellschaftliche Praxis einzelne Personen statistischen Gruppen zuordnet, werden die vorausgesagten statistischen Eigenschaften auf diese konkreten Personen angewendet. Die so entstehenden Missbrauchspotenziale würden vom geltenden Datenschutzrecht nicht reguliert und die Verwendung anonymisierter Massendaten finde in einem weitestgehend rechtsfreien Raum statt. Mühlhoff plädiert deswegen für einen datenschützerischen Ansatz, bei dem einerseits prädiktive Informationen rechtlich personenbezogenen Daten gleichgestellt werden und andererseits in definierten Anwendungsbereichen (z. B. bei Haftentscheidungen) die Herstellung prädiktiver Risiko-Modelle untersagt wird.

*Rita Jordan* geht in ihrem Kapitel ebenfalls von der Beobachtung aus, dass mit dem Einsatz selbstlernender Algorithmen nicht nur der Umfang und die Geschwindigkeit, mit der Daten erfasst, verarbeitet und ausgewertet werden, zunimmt, sondern auch die Abgrenzbarkeit zwischen personenbezogenen und nicht personenbezogenen Daten verschwimmt. Da-

durch gerieten die Zwecke des Datenschutzrechts (Persönlichkeitsschutz, informationelle und demokratische Selbstbestimmung) und seine Schutzprinzipien (u. a. Zweckbindung, Datenminimierung und Transparenz) in Spannung zu den Gewinninteressen datenbasierter Geschäftsmodelle und dem herrschenden Innovationsdruck. Die Abgrenzbarkeit von personenbezogenen und nicht-personenbezogenen Daten sei aber zentral für das dogmatische Fundament der EU-Datenschutz-Grundverordnung. Jordan macht deutlich, wie individuellen Nutzerinnen und Nutzern eine aufgeklärte Rechtsausübung praktisch erschwert wird, beispielsweise durch immer kleinteiligere Datenschutzerklärungen. Sie erläutert, wie sich dies bei der Digitalisierung von Städten manifestiert, wo sich die Innovationskraft algorithmischer Datenverarbeitung für Nachhaltigkeits- und Verkehrsziele mit der physischen Oberfläche urbaner Erfahrungs- und Handlungsräume verschränken soll. Wegen der Ubiquität der erfassten Daten und der damit einhergehenden Risiken für Privatheit und Selbstbestimmung sei eine grundlegende Rekonzeptualisierung des Datenschutzrechts sowie eine Demokratisierung der Technologieentwicklung –- insbesondere im Bereich KI-basierter Technologien – in städtischen Räumen notwendig.

*Jörn Lamla* (Universität Kassel) beleuchtet in seinem Kapitel über die KI als hybride Lebensform schließlich das Wechselverhältnis von Mensch und digitaler Anwendung: KI setze mit ihren Herausforderungen das humanistische Selbstverständnis unter Druck. Der Beitrag argumentiert, dass dies zurecht geschieht, dabei jedoch mit einer verkürzenden Gegenüberstellung operiert wird. Demnach seien KI-Technologien zwar paradigmatisch für die expansive Dynamik hybrider Lebensformen, die Menschen und Maschinen in Feedbackschleifen verklammern, deren Charakter werde aber immer noch verkannt. Die Technologie entwickle sich zu einem Paradigma, das nach Lamla drei Aspekte umfasst, die bei der Analyse des Verhältnisses von Mensch und Maschine und der gesellschaftlichen Auswirkungen zusammen gedacht werden müssten: 1) die sich verstärkende Hybridisierung von Mensch und Maschine, 2) die Datafizierung des Lebens und 3) eine Algorithmisierung, also eine permanente Weiterentwicklung und das Lernen von Algorithmen aus Hybridisierung und Datafizierung. Angesichts der zentralen Rolle, die Digitalisierung und insbesondere KI-Technologien in unserer Gesellschaft spielen, plädiert Lamla entgegen der vorherrschenden kybernetischen Sichtweise für eine Reflektion der Dominanzstruktur des digitalen Analogismus. Um dieser Entwicklung wirksam und kritisch entgegenzutreten, so die These, braucht es mehr als die Beschwörung humanistischer Werte: Es bedürfe eines besseren Verständnisses für die ontologische Heterogenität der gesellschaftlichen Existenzweisen, die in hybriden Lebensformen versammelt sind.

*Künstliche Intelligenz, Profiling und Überwachung*

Überwachung und Profiling (vor allem für staatliche Akteure wie Strafverfolgungsbehörden und Geheimdienste) sind seit langem treibende Kräfte bei der Entwicklung von KI-Verfahren. Die Beiträge in Teil II fokussieren auf die Fragen, welche Rolle KI hier spielen kann, wie effektiv Betroffenenrechte gewährleistet werden können und wie gut das entstehende europäische Recht auf die absehbaren Herausforderungen reagiert.

*Stephan Schindler* und *Sabrina Schomberg* (Universität Kassel) beleuchten den aktuellen Verordnungsentwurf der Europäischen Kommission zur Regulierung künstlicher Intelligenz (AI Act), mit dem ein einheitlicher Rechtsrahmen für die Entwicklung, Vermarktung und Verwendung künstlicher Intelligenz im Einklang mit den Werten der Europäischen Union geschaffen werden soll. Sie stellen dabei die Frage, ob es sich mit Blick auf Anwendungen der biometrischen Erkennung um einen großen Wurf oder lediglich um Symbolpolitik handelt. Die biometrische Erkennung nimmt im Verordnungsentwurf eine herausgehobene Stellung ein; insbesondere sieht sie ein Verbot der Verwendung biometrischer Echtzeit-Fernidentifizierungssysteme in öffentlich zugänglichen Räumen zu Strafverfolgungszwecken vor. Von diesem Verbot gäbe es allerdings zahlreiche Ausnahmen, so dass die biometrische Echtzeit-Fernidentifizierung in vielen spezifischen Anwendungskontexten mit mehr oder weniger strikten Auflagen (Dokumentations- und Aufzeichnungspflichten, menschliche Aufsicht) doch betrieben werden könne. Insgesamt begrüßen Schomberg und Schindler den Verordnungsentwurf, kritisieren aber die Ausnahme in ihrer Vielzahl und Breite als problematisch und weisen darüber hinaus auf weitere offene Fragen hin, die insbesondere den Einsatz biometrischer Systeme durch staatliche Stellen zu Strafverfolgungszwecken betreffen.

*Jasmin Schreyer* (Universität Erlangen-Nürnberg) untersucht in ihrem Kapitel den Datenschutz als zentrale Machtfrage in der Plattformökonomie. Spätestens seit den Snowden-Enthüllungen sei klar, dass das Internet mit seinen scheinbar unbegrenzten Möglichkeiten zur Datensammlung ein Herrschaftsinstrument sei, das nicht nur von staatlichen Akteuren, sondern vor allem auch von international agierenden Datenunternehmen genutzt wird. Obwohl die früheren Hoffnungen auf eine demokratisierende Wirkung des Internet mittlerweile ad absurdum geführt worden seien, inszenierten sich die Plattformanbieter als neutrale Vermittlungsinstanzen und propagierten, dass ihre Datensammlungen eine Form der „höheren" Intelligenz ermögliche, die Wissen, Wahrheit und Objektivität generiere. Schreyer zeigt auf, welche Wirkung das von den Akteuren akkumulierte Wissen über vergangene, gegenwärtige und zukünftige Präferenzen, Ein-

stellungen und Verhalten auf die betroffenen Subjekte hat. Dies führe bei den betroffenen Subjekten zu einer Internalisierung des Machtverhältnisses sowie zu einer Selbstkontrolle und Normierung des Verhaltens. Die Autorin betont, dass sich dieser panoptische Zustand weiter verschärfen werde.

*Matthias Marx* und *Alan Dahi* berichten in ihrem Kapitel über praktische Erfahrungen bei der Durchsetzung von Betroffenenrechten beim US-amerikanischen Unternehmen Clearview AI, das sich auf KI-gestützte Gesichtserkennung spezialisiert hat. Im Jahr 2020 wurde bekannt, dass Clearview AI zum Zwecke der Gesichtserkennung rechtswidrig mehr als zwanzig Milliarden Fotos von Gesichtern im Internet gesammelt und ausgewertet hatte. Die Autoren zeichnen den Weg einer Beschwerde samt der dabei auftretenden Hindernisse nach, die beim Hamburgischen Beauftragten für Datenschutz und Informationsfreiheit eingereicht wurde. Zudem beleuchten sie einige der rechtlichen Fragen, darunter die Anwendbarkeit der DSGVO, die Rechtmäßigkeit der Verarbeitung sowie die Handlungsmöglichkeiten der Aufsichtsbehörden. Schließlich werden Entscheidungen anderer europäischer Aufsichtsbehörden zu Clearview AI kurz vorgestellt. Der Beitrag demonstriert, wie schwierig die Wahrnehmung grundlegender Betroffenenrechte im Falle eines US-amerikanischen Unternehmens sein kann.

Schließlich befassen sich *Marianne von Blomberg* und *Hannah Klöber* (Universität Köln) in ihrem Beitrag mit dem chinesischen Sozialkreditsystem (SKS), das nicht nur die finanzielle Kreditwürdigkeit der Bürger, sondern deren Vertrauenswürdigkeit im weiteren Sinne ermitteln soll. Die Pläne sehen vor, dass Sozialkreditdossiers für natürliche Personen auf zentraler Ebene angelegt und darin Informationen über ordnungs- und gesetzeswidriges Verhalten gespeichert werden. Anders als ihre Vorgänger sollen die modernen Sozialkreditdossiers transparent, den betroffenen Personen zugänglich und von ihnen korrigierbar sein. Der Beitrag beleuchtet deshalb die lange Tradition personenbezogener Dossiers in China und fragt, ob sich das SKS fundamental von vorherigen Dossiersystemen unterscheidet. Dazu analysieren die Autorinnen den aktuellen Rechtsrahmen für personenbezogene Sozialkreditdossiers im Hinblick auf den Transparenzanspruch des SKS. Sie erläutern, dass eine wachsende Anzahl von lokalen und sektoralen Verordnungen die Verwaltung persönlicher Sozialkreditinformationen regulieren. Ihre Vielfältigkeit einerseits und die nicht standardisierte Sammlung und Verarbeitung von Informationen unter Einbeziehung verschiedener Akteure andererseits erschwerten jedoch das Einsehen und die Korrektur der Dossiers. Um dem Anspruch der Transparenz gerecht zu werden bedürfte es daher einer Vereinheitlichung

des rechtlichen Rahmens des SKS und einer eindeutigen Definition von „Sozialkredit".

*Künstliche Intelligenz und Nutzendenverhalten*

Die Beiträge in Teil III befassen sich mit der Frage, welche menschlichen Faktoren bei der Wahrung von Privatheit und Selbstbestimmung eine Rolle spielen. Dazu werden einerseits KI-basierte Möglichkeiten diskutiert, die typische menschliche Faktoren entweder ausnutzen oder die Nutzenden bei einem Datenschutz wahrenden Verhalten unterstützen können. Andererseits werden menschliche Faktoren im Umgang mit KI am Beispiel der Nutzung von Messenger-Diensten diskutiert.

Der Beitrag von *Hannah Ruschemeier* (Fernuniversität Hagen) dreht sich um das so genannte *Privacy Paradox*, welches beschreibt, dass Menschen zwar regelmäßig bekunden, wie wichtig ihnen Privatsphäre und Datenschutz ist, dieser Selbsteinschätzung aber keine entsprechenden Taten folgen lassen. Unternehmen nutzten dieses Phänomen aus oder förderten es sogar, so dass viele Personen trotz der betonten Wichtigkeit von Privatheit und Selbstbestimmung niedrigschwellig oder gar anlasslos persönlichste Informationen über sich preisgeben. Diese Diskrepanz zwischen Selbsteinschätzung und realem Verhalten sollte – so die Argumentation der Autorin – vom Recht nicht unbeachtet bleiben. Privatheit als Konzept in der Vorstellung vieler Menschen könne unendlich viele Facetten abdecken, die sich nur teilweise oder auch gar nicht mit konkreten persönlichen Verhaltensweisen überschneiden. Das Recht reflektiere diese realen Voraussetzungen von Privatheit jedoch bisher unzureichend, wie das Beispiel der datenschutzrechtlichen Einwilligung zeige. Zur Adressierung dieser Problemlage wird eine veränderte Ausrichtung des Datenschutzes von einem höchstpersönlichen Gut hin zur Regelung kollektiver Auswirkungen und institutioneller Verantwortung angeregt.

*Leen Al Kallaa* und Kolleginnen und Kollegen (Universität Bochum) befassen sich in ihrem Kapitel mit der Rolle, die Datenschutz und Datensicherheit bei der Messenger-Auswahl und -Nutzung unter arabischsprachigen Nutzerinnen und Nutzer spielen. Wie bei anderen Nutzendengruppen gehörten Instant Messenger auch bei dieser Gruppe, die in anderen Untersuchungen meist unterrepräsentiert ist, zu den am häufigsten genutzten Smartphone-Apps. Im Rahmen einer empirischen Untersuchung fand das Autorenteam heraus, dass die Änderung wichtiger Datenschutzaspekte in den Nutzungbedingungen von Whatsapp im Frühjahr 2021 von der befragten Gruppe überwiegend nicht wahrgenommen wurde: Lediglich 8 %

der Befragten hätten einen Messenger-Wechsel erwogen. Insgesamt bestätigt die Studie, dass die Gründe gegen den Wechsel zu einem sichereren Messenger vor allem die Netzwerkeffekte sind: An erster Stelle steht die Frage, wie viele Bekannte man erreichen kann.

### *Künstliche Intelligenz, Desinformation und Deepfakes*

Teil IV dreht sich um Fragen der Desinformation, zu deren Erstellung und Verbreitung seit einigen Jahren erfolgreich KI-Verfahren genutzt werden. Dies reicht von der Extraktion von Persönlichkeitsmerkmalen, über Social Bots und Verfahren des Mikrotargeting bis hin zu Deepfakes, also realistisch wirkende, aber synthetische Medieninhalte. Während bspw. der Einsatz von Mikrotargeting im US-Präsidentschaftswahl 2012 noch als modern und innovativ galt, wurde spätestens mit dem Fall „Cambridge Analytica" klar, welches Gefahrenpotenzial hier für die demokratischen Strukturen und Prozesse sowie deren Standards entsteht. Seither sind Bestrebungen im Gange die Gefahren mit unterschiedlichsten Mittel einzuhegen.

Zunächst widmen sich *Anna Louban* (HWR Berlin) und Kolleginnen und Kollegen dem relativ neuen Phänomen der Deepfakes, also durch KI-Methoden generierte oder manipulierte Bilder, Audios und Videos, die politische Desinformation und Propaganda in videographischer Form transportieren können. Sie fragen interdisziplinär aus den Perspektiven der Rechts- und Politikwissenschaft sowie der Informatik nach den Risiken für politische Entscheidungsprozesse, zu denen Deepfakes und ihre Nutzung für politische Desinformation führen können. Darauf basierend präsentiert der Beitrag Ansätze aus dem multidisziplinär ausgerichteten Forschungsprojekt FAKE-ID zur Erforschung KI-basierter Deepfake-Detektoren.

*Lena Isabell Löber* (Universität Kassel) untersucht die Möglichkeiten, die die KI bietet, um Dienstbetreiber bei der Erfüllung der gesetzlichen Pflichten zur Bekämpfung von Hasskriminalität im Netz zu unterstützen. KI-Lösungen können wirkungsvolle Instrumente sein, um schädlichen Inhalte und Manipulationstechniken wie Social Bots in sozialen Medien zu detektieren. Die mit ihrem Einsatz verbundenen Risiken für Kommunikationsgrundrechte und Meinungspluralität müssen aber durch manuelle Nachkontrollen automatisiert ermittelter Treffer und einen verfahrensorientierten Grundrechtsschutz eingehegt werden. Außerdem hält die Autorin schärfere Transparenzvorgaben und Aufsichtsstrukturen für erforderlich, um den Risiken der technisch-organisatorischen Gestaltungs- und

Entscheidungsmacht großer Anbieter von sozialen Netzwerken z. B. im Rahmen der algorithmischen Empfehlungssysteme zu begegnen. Betrachtet werden zu diesem Zweck die neuen Regelungen im Medienstaatsvertrag und Netzwerkdurchsetzungsgesetz, die zu mehr Transparenz für die Betroffenen führen sollten, aber gerade beim Themenkomplex Desinformation weitestgehend vage bleiben. Dem gegenübergestellt werden die auf EU-Ebene im Rahmen der Entwürfe für die KI-Verordnung und den Digital Services Act vorgesehenen Regelungen, die auch weitergehende Pflichten vorsehen und einen wichtigen Beitrag zu einem ganzheitlichen Ansatz im Umgang mit digitaler Desinformation leisten könnten.

Das Kapitel von *Nicole Krämer* (Universität Duisburg-Essen) und Kolleginnen und Kollegen diskutiert schließlich aus interdisziplinärer Perspektive die Probleme von Desinformation über Messengerdienste. Aus Sicht der Informatik, Journalistik, Medienpsychologie und Rechtswissenschaften werden jeweils der Stand der Forschung zur Fragestellung und zur Lösung durch denkbare Werkzeuge dargestellt, eigene Ansätze und Beiträge diskutiert und Fragestellungen herausgearbeitet, die als Grundlage für eine gemeinsame Forschung dienen können. So entsteht ein Überblick über die zahlreichen Perspektiven, mit denen an die Thematik herangegangen werden kann. Basierend darauf werden exemplarisch die Einflüsse datenschutzrechtlicher Projektentscheidungen auf die Projektarbeit diskutiert.

### *Einsatz von KI in Gesundheit und Pflege*

Im abschließenden Teil V des Bandes werden in zwei Kapiteln Beispiele des Einsatzes von KI im Bereich von Gesundheit und Pflege genauer beleuchtet, also aus einem Bereich, wo sowohl die Erwartung an das Gemeinwohl aber auch die potenziellen Risiken für den Einzelnen am höchsten sind.

*Roger von Laufenberg* (Wiener Zentrum für sozialwissenschaftliche Sicherheitsforschung) betrachtet KI-Systeme in Pflegeeinrichtungen für ältere Menschen. Die Technisierung der Pflege sei vor allem eine Reaktion auf die alternde Bevölkerung und der damit einhergehenden Pflegekrise. Während dies in der Theorie durchaus erfolgversprechend scheint, beschreibt der Beitrag anhand einem Fallbeispiels (Sturzdetektion), dass die Entwicklung von KI-Pflegetechnologien häufig von der alltäglichen Lebensrealität älterer Personen entkoppelt ist. Dabei wird einerseits deutlich, wie in den unterschiedlichen Schritten in der Systementwicklung ein Bild von älteren Personen gezeichnet wird, das von Vulnerabilität geprägt ist. Andererseits erhielten ältere Personen als direkt Betroffene keine Möglichkeit, ihre

Sichtweisen in die Entwicklung und Implementierung mit einzubringen. Dadurch entstünden KI-Systeme, die den Anspruch von Fürsorge für ältere Menschen haben, dazu aber auf umfassende Überwachung ausgelegt sind und mögliche Risiken und negative Auswirkungen für Privatheit und Selbstbestimmung häufig ausblenden.

Im abschließenden Kapitel analysieren *Niël H. Conradie* (RWTH Aachen) und Kolleginnen und Kollegen, welche Auswirkungen intelligente Wearables – mit Bio-Sensoren ausgestattete kleine Computersysteme, die direkt am Körper getragen werden – auf die Entscheidungsfreiheit von schutzbedürftigen Personen haben. Der Markt für Wearables boomt seit einige Jahren und ist immer noch ein weitgehend unreguliertes Experimentierfeld für mehr oder weniger sinnvolle Anwendungen. Wie bei den meisten neu aufkommenden Technologien müssen die Vorteile und Risiken bewertet und gegeneinander abgewogen werden. Besonders wichtig ist diese Abwägung, wenn es sich um Anwendungen handelt, die schutzbedürftige Personengruppen betreffen, da diese oft und in besonderem Maße von Verletzungen der Selbstbestimmung betroffen sind. Dieser Beitrag untersucht aus einer explizit normativen und ethischen Perspektive die potenziellen Auswirkungen von Smart Wearables auf die Autonomie der Entscheidungsfindung in drei solchen Gruppen, nämlich: Kinder, ältere Erwachsene und Personen mit nicht altersbedingten Autonomieeinschränkungen.

# Teil I
# Künstliche Intelligenz und Selbstbestimmung

# Prädiktive Privatheit: Kollektiver Datenschutz im Kontext von Big Data und KI

*Rainer Mühlhoff*

Zusammenfassung

Big Data und künstliche Intelligenz (KI) stellen eine neue Herausforderung für den Datenschutz dar. Denn diese Techniken werden dazu verwendet, anhand der anonymen Daten vieler Menschen Vorhersagen über Dritte zu treffen – etwa über Kaufkraft, Geschlecht, Alter, sexuelle Orientierung, ethnische Zugehörigkeit, den Verlauf einer Krankheit etc. Die Grundlage für solche Anwendungen „prädiktiver Analytik" ist ein Vergleich von Verhaltensdaten (z.B. Nutzungs-, Tracking- oder Aktivitätsdaten) des betreffenden Individuums mit den potenziell anonymisiert verarbeiteten Daten vieler Anderer anhand von Machine Learning Modellen oder einfacherer statistischer Verfahren. Der Artikel weist zunächst darauf hin, dass mit prädiktiver Analytik erhebliche Missbrauchspotenziale verbunden sind, welche sich als soziale Ungleichheit, Diskriminierung und Ausgrenzung manifestieren. Diese Missbrauchspotenziale werden vom geltenden Datenschutzrecht (EU DSGVO) nicht reguliert; tatsächlich findet die Verwendung anonymisierter Massendaten in einem weitestgehend rechtsfreien Raum statt. Unter dem Begriff "prädiktive Privatheit" wird ein datenschützerischer Ansatz vorgestellt, der den Missbrauchsgefahren prädiktiver Analytik begegnet. Die prädiktive Privatsphäre einer Person oder Gruppe wird verletzt, wenn anhand der Daten vieler anderer Individuen ohne ihr Wissen und gegen ihren Willen sensible Informationen über sie vorausgesagt werden. Prädiktive Privatheit wird sodann als Schutzgut eines kollektiven Ansatzes im Datenschutz formuliert und verschiedene Verbesserungen der DSGVO im Hinblick auf die Regulierung prädiktiver Analytik werden vorgeschlagen.

## *1. Einleitung*

Eine der aktuell wichtigsten Anwendungen von KI-Technologie ist die sogenannte prädiktive Analytik. Unter diesen Begriff fasse ich datenbasierte Vorhersagemodelle, die über beliebige Individuen anhand verfügbarer

Daten Prognosen stellen. Diese Prognosen können sich auf zukünftiges Verhalten beziehen (z.B., was wird jemand wahrscheinlich kaufen?), auf unbekannte persönliche Attribute (z.B. sexuelle Identität, ethnische Zugehörigkeit, Wohlstand, Bildungsgrad) oder auf persönliche Risikofaktoren (z.B. psychische oder körperliche Krankheitsdispositionen, Suchtverhalten oder Kreditrisiko). Prädiktive Analytik ist brisant, denn neben den gesellschaftlich nutzbringenden Anwendungsmöglichkeiten besitzt die Technologie ein enormes Missbrauchspotenzial und ist aktuell gesetzlich kaum reguliert. Prädiktive Analytik ermöglicht die automatisierte und daher großflächige Ungleichbehandlung von Individuen und Gruppen beim Zugriff auf ökonomische und gesellschaftliche Ressourcen wie Arbeit, Bildung, Wissen, Gesundheitsversorgung und Rechtsdurchsetzung. Speziell im Kontext von Datenschutz und Antidiskriminierung muss die Anwendung prädiktiver KI-Modelle als eine neue Form von Datenmacht großer IT-Unternehmen analysiert werden, die im Zusammenhang mit der Stabilisierung und Hervorbringung von Strukturen der Diskriminierung, der sozialen Klassifizierung und der datenbasierten sozialen Ungleichheit steht.

Vor dem Hintergrund der enormen gesellschaftlichen Auswirkungen prädiktiver Analytik werde ich in diesem Kapitel argumentieren, dass wir im Kontext von Big Data und KI neue Ansätze im Datenschutz benötigen. Mit dem Begriff *prädiktive Privatheit* werde ich den Schutz der Privatheit einer Person oder Gruppe gegen ihre neuartige Form der Verletzbarkeit durch *abgeleitete* oder *vorhergesagte* Informationen fassen und normativ verankern. Die Anwendung prädiktiver Modelle auf Einzelindividuen, um damit Entscheidungen zu stützen, stellt eine Verletzung der Privatheit dar – die jedoch neuartigerweise weder durch „Datenklau" noch durch einen Bruch von Anonymisierung zustande kommt. Die Verletzung der prädiktiven Privatheit erfolgt mittels eines Abgleichs der über das Zielindividuum bekannten Hilfsdaten (z.B. Nutzungsdaten auf Social Media, Browserverlauf, Geo-Location-Daten) mit den Daten vieler tausend *anderer* Nutzer:innen. Prädiktive Analytik verfährt nach dem Prinzip des „pattern matching" und ist immer dort möglich, wo es eine hinreichend große Gruppe von Nutzer:innen gibt, welche die sensiblen Zielattribute über sich preisgibt, weil sie sich der Big Data-basierten Verwertungsweisen nicht bewusst sind oder denkt, „nichts zu verbergen zu haben". Deshalb markiert das Problem der prädiktiven Privatheit eine Grenze des im Datenschutz weit verbreiteten Individualismus und gibt dazu Anlass, kollektivistische Schutzgüter und kollektivistische Abwehrrechte im Datenschutz zu verankern.

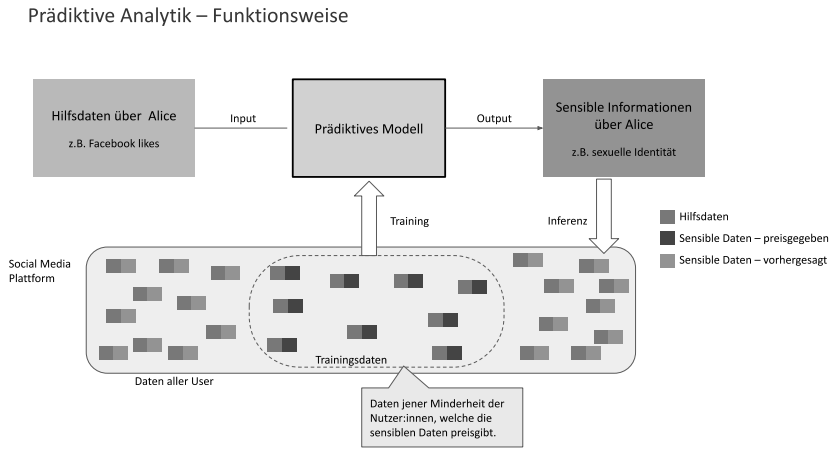
Eine solche kollektivistische Perspektive im Datenschutz berücksichtigt erstens, dass Individuen *nicht* in jeder Hinsicht frei entscheiden können sollten, welche Daten sie über sich gegenüber modernen Datenunternehmen preisgeben, denn die eigenen Daten können potenziell negative Auswirkungen auch auf andere Individuen haben. Zweitens bringt diese kollektivistische Perspektive zur Geltung, dass große Ansammlungen anonymisierter Daten vielen Individuen aufgrund der darin „lernbaren" Korrelationen zwischen sensiblen und weniger sensiblen Datenfeldern von Datenverarbeitenden *nicht* frei verarbeitet werden können sollten, wie es die aktuelle Rechtslage nach DSGVO bei anonymisierten Daten erlaubt. Drittens schließlich werde ich fordern, dass die Betroffenenrechte des Datenschutzes (Recht auf Auskunft, Rektifizierung, Löschung, etc.) kollektivistisch neu formuliert werden sollten, so dass betroffene Kollektive und das Gemeinwesen im Ganzen befugt wären, solche Rechte im Sinne des Gemeinwohls gegenüber datenverarbeitenden Organisationen auszuüben.

## *2. Prädiktive Analytik*

Für den Gegenstand dieses Artikels ist es unerheblich, auf welchen Algorithmen und Verfahren ein prädiktives Modell konkret beruht. Es handelt sich bei prädiktiver Analytik um einen Container-Begriff, der sowohl Verfahren des maschinellen Lernens als auch einfachere statistische Auswertungen umfasst. Während prädiktive Analytik die technologische Disziplin bezeichnet, verweist „prädiktives Modell auf eine konkrete Manifestation dieser Technologie. Jedoch ist für ein adäquates Verständnis des Datenschutzproblems eine funktionale Charakterisierung prädiktiver Modelle hilfreich. Es handelt sich dabei um Datenverarbeitungssysteme, die als Input eine Reihe verfügbarer Daten über ein Individuum (oder einen „Fall") erhalten und als Ausgabe die Schätzung einer unbekannten Information, eine Klassifikation oder eine Entscheidung in Bezug auf das Individuum angeben (im Folgenden kurz „Zielvariable" genannt).

Die Inputdaten sind dabei typischerweise leicht verfügbare Hilfsdaten, zum Beispiel Trackingdaten, der Browser- oder Standort-Verlauf, oder Social Media Daten (Likes, Postings, Freund:innen, Gruppenmitgliedschaften). Bei der Zielvariablen handelt es sich typischerweise um schwer zugängliche oder besonders sensible Informationen über das Individuum, oder um eine Entscheidung über das Individuum in Bezug auf die Geschäftsvorgänge des Betreibers des prädiktiven Modells (zum Beispiel: zu welchem Preis dem Individuum eine Versicherung oder ein Kredit angeboten wird).

In der prädiktiven Analytik möchte man also anhand leicht zugänglicher Daten schwer zugängliche Informationen über Individuen abschätzen. Dazu vergleichen prädiktive Modelle den durch die Inputdaten gegebenen Fall nach Prinzipien der Mustererkennung mit Tausenden oder Millionen anderen Fällen, die das Modell zuvor während einer Lernphase (oder mittels anderer, statistischer Verfahren) ausgewertet hat. Häufig werden solche Modelle mit Verfahren des überwachten Lernens trainiert. Dazu wird eine große Menge Trainingsdaten benötigt, also ein Datensatz, in dem für eine Kohorte von Individuen beide Datenfelder, die Hilfsdaten und die Zieldaten, erfasst sind. Solche Datensätze fallen regelmäßig im Kontext sozialer Alltagsmedien an, zum Beispiel produziert die Teilmenge aller Facebook-Nutzer:innen, die in ihrem Profil explizit Angaben über ihre sexuelle Orientierung machen, einen Trainingsdatensatz für prädiktive Modelle zur Abschätzung der sexuellen Orientierung *beliebiger* Facebook-Nutzer:innen anhand der auf Facebook anfallenden Nutzungsdaten, wie zum Beispiel Facebook-Likes (s. Abb. 1).




*Schematische Darstellung der Vorgehensweise prädiktiver Analytik*

Wenn nur wenige Prozent der mehr als zwei Mrd. Facebook-Nutzer:innen Angaben über ihre sexuelle Orientierung machen, dann sind das einige Millionen Nutzer:innen. Das damit trainierbare prädiktive Modell kann die Plattform im nächsten Schritt dazu verwenden, die sexuelle Orientierung für alle anderen Facebook-Nutzer:innen abzuschätzen – auch für Nutzer:innen, die der Verarbeitung dieser Information nicht zustimmen

würden, diese Angabe bewusst nicht getätigt haben oder möglicherweise nicht wissen, dass das Unternehmen in der Lage ist, diese Informationen über sie abzuschätzen (vgl. auch Skeba und Baumer 2020).

Mediziner:innen von der University of Pennsylvania haben gezeigt, dass sich mit dieser Vorgehensweise anhand von Facebook-Daten beispielsweise vorhersagen lässt, ob eine Nutzer:in an Krankheiten wie Depression, Psychosen, Diabetes oder Bluthochdruck leidet (Merchant u.a. 2019). Facebook selbst hat bekannt gegeben, suizidale Nutzer:innen anhand ihrer Postings erkennen zu können (Goggin 2019). Eine viel beachtete Studie von Kosinski et al. zeigt, dass die Daten über Facebook-Likes dazu verwendet werden können, „eine Reihe höchst sensibler persönlicher Attribute vorherzusagen, darunter sexuelle Orientierung, Ethnie, religiöse und politische Ansichten, Persönlichkeitseigenschaften, Intelligenz, happiness, Suchtverhalten, Trennung der Eltern, Alter und Geschlecht“ (Kosinski u.a. 2013).

Solche prädiktiven Analysen stoßen bei Versicherungs- und Finanzkonzernen auf großes Interesse, weil sie eine individuelle Risikobemessung jenseits der klassischen Credit Scores erlauben.[1] Auch im Personalmanagement werden solche prädiktiven Modelle verwendet, um zum Beispiel eine automatisierte Vorauswahl von Bewerber:innen bei Einstellungsvorgängen durchzuführen (O’Neil 2016, S. 108, 148). Zu den ersten und häufigsten Anwendungen prädiktiver Analytik gehört außerdem die gezielte Werbung (targeted advertising). So ist es einer US-amerikanischen Supermarktkette im Jahr 2011 gelungen, anhand der Einkaufsdaten, die über Rabattprogramme (customer loyalty cards) gesammelt werden, schwangere Kundinnen zu identifizieren (Duhigg 2012).

## 3. *Prädiktive Privatheit*

Prädiktive Analytik erlaubt es, unbekannte oder potenziell sensible Informationen über Individuen oder Gruppen anhand vermeintlich weniger sensibler und leicht verfügbarer Daten (Hilfsdaten) abzuschätzen. Dies ist mit modernen maschinellen Lernverfahren möglich, wenn viele Nutzer:innen einer digitalen Plattform die Datengrundlage schaffen, um Korrelationen zwischen den Hilfsdaten und den Zielinformationen zu ermitteln. Wir stehen hier also vor einer Situation, in der die *Datenfreigiebigkeit*

1 Siehe Lippert 2014 zum Beispiele der Firma ZestFinance sowie O’Neil 2016, Kap. 8 zu sogenannten “e-scores” als alternative credit scoring-Verfahren.

*einer Minderheit von Nutzer:innen* (zum Beispiel die prozentual wenigen Facebook-Nutzer:innen, die Angaben über ihre sexuelle Orientierung machen) den Standard der über *alle* Gesellschaftsmitglieder ableitbaren Informationen setzt. In der industriellen Verwendung prädiktiver Analytik im Kontext digital vernetzter Medien hat sich eine Praxis etabliert, in der die von Vorhersagen betroffenen Individuen in den meisten Fällen nicht informiert oder gefragt werden. Auch auf regulatorischer Ebene ist das Problem bisher im EU-Kontext weitestgehend unbeleuchtet: Insbesondere die DSGVO verfehlt es, die Herstellung oder Verwendung prädiktiver Modelle an geeignete Voraussetzungen zu knüpfen oder verantwortungsvoll einzuschränken.[2]

Vorhergesagte Informationen über Individuen oder Gruppen ermöglichen neben vorstellbar nutzbringenden Anwendungen zahlreiche schädliche und missbräuchliche Verwendungsweisen, welche mit Diskriminierung, Ungleichbehandlung und weiteren Grundrechtseingriffen der Betroffenen verbunden sein können. Um einen Schutz vor der missbräuchlichen Verwendung abgeschätzter Informationen normativ zu verankern – zunächst ethisch, sodann politisch und rechtlich –, möchte ich deshalb ein neues Schutzgut konstruieren. In direkter Antwort auf die Gefahrenlage der prädiktiven Analytik schlage ich dazu den Begriff der *prädiktiven Privatheit* vor (vgl. Mühlhoff 2020b, Mühlhoff 2021).[3] Prädiktive Privatheit lässt sich am besten negativ definieren, indem fixiert wird, wann sie *verletzt* ist:

> Die prädiktive Privatheit einer Person oder Gruppe wird verletzt, wenn sensible Informationen ohne ihr Wissen oder gegen ihren Willen über sie vorhergesagt werden, und zwar in solcher Weise, dass daraus die Ungleichbehandlung eines Individuums oder einer Gruppe resultieren könnte. (vgl. Mühlhoff 2021)

---

2 In diesem Sinne argumentiert auch Roßnagel (2018, S. 365–367) für eine Modernisierung der DSGVO angesichts der Gefahr durch prognostizierte Informationen.

3 Es gibt verwandte Begriffsvorschläge, die in eine ähnliche Richtung zielen. Darunter ist insbesondere „categorical privacy" von Vedder 1999 zu erwähnen, sowie die jüngere Debatte zu „group privacy" im Kontext von Big Data (Floridi 2014; Taylor u.a. 2016; Mittelstadt 2017) und „inferential privacy" (Loi und Christen 2020). Auch die Arbeiten zu einem „right to reasonable inferences" von Sandra Wachter und Brent Mittelstadt (Wachter und Mittelstadt 2018) schlagen eine ähnliche Richtung ein. Eine Auseinandersetzung mit den Gemeinsamkeiten und Unterschieden dieser Begriffe zu dem hier konstruierten Konzept der prädiktiven Privatheit findet sich in Mühlhoff 2021.

Hinter dieser sehr allgemeinen Begriffsbildung steht zunächst das Anliegen, angesichts der durch KI veränderten technologischen Situation auch die gesellschaftliche und kulturelle Auffassung von Privatheit anzupassen und zu erweitern. Denn bisher hat man sich unter Verletzungen von (informationeller) Privatheit meist einen nicht-autorisierten Zugriff auf die private „Informationssphäre" oder Eingriffe in die informationelle Selbstbestimmung des Einzelnen vorgestellt, durch die dem Datensubjekt Informationen „entwendet" werden, die es nicht über sich preisgeben wollte.[4] Zwar werden bei einer Verletzung prädiktiver Privatheit ebenfalls Informationen gewonnen, die das betroffene Subjekt mutmaßlich nicht preisgeben möchte, jedoch geschieht dies nicht auf dem Weg der „Entwendung" oder des Eindringens in eine private Sphäre (diese Metapher ist in der neuen technologischen Situation längst nicht mehr adäquat, siehe dazu auch Ruschemeier in diesem Band). Vielmehr werden die Informationen über das Datensubjekt abgeschätzt, und zwar anhand eines Vergleichs mit den Daten, die viele *andere* Datensubjekte über sich preisgeben. Hierbei kommt es darauf an, dass diese Verletzungen prädiktiver Privatheit *nicht* von der Genauigkeit oder Korrektheit der geschätzten Informationen abhängen, sondern allein davon, dass diese Informationen das Potenzial einer Ungleichbehandlung der betroffenen Individuen oder Gruppen bergen. Das heißt, es wäre nach der ethischen und datenschützerischen Norm der prädiktiven Privatheit nicht automatisch legitim, Menschen anhand von über sie vorhergesagten Informationen unterschiedlich zu behandeln, bloß weil die Vorhersagen bestimmte Anforderungen der Genauigkeit erfüllt.[5]

---

4 Zum Begriff Privatheit werden häufig zwei oder mehr Haupttraditionen unterschieden, die im anglophonen Raum als „nonintrusion theory" und als „control theory" of privacy in Erscheinung treten (vgl. Tavani 2007, der insgesamt vier Kategorien unterscheidet). Das Verständnis von Privatheit als Nicht-Intrusion betont dabei eine (oder sogar mehrere geschachtelte) private Sphäre(n) jedes Individuums, die vor Einblicken und Eingriffen zu schützen sei(en); Kontrolltheorien setzen dagegen weniger auf die Abgeschiedenheit für sich, sondern auf das Vermögen des Individuums, effektiv und potenziell differenziert darüber zu verfügen, wer welchen „Zugang" zu den eigenen persönlichen Informationen hat.

5 In diesem Punkt weicht die ethische und datenschützerische Norm der prädiktiven Privatheit von der zu kurz greifenden Forderung eines „right to reasonable inferences" von Sandra Wachter und Brent Mittelstadt (vgl. Wachter und Mittelstadt 2018) ab.

## 4. Ein neues Datenschutzproblem: Drei Angriffstypen

Die Abschätzung persönlicher und potenziell sogar sensibler Informationen über Individuen anhand von Massendaten stellt ein neues dominantes Angriffsszenario im Datenschutz unter den Bedingungen unzureichend regulierter KI- und Big Data-Technologie dar. Dies ist ein Angriffsszenario, das erst seit etwa zehn Jahren virulent ist. Um die neue Qualität dieser Herausforderung und die entsprechend neuen Schutzbedarfe herauszuarbeiten, lohnt sich eine vergleichende Gegenüberstellung des neuartigen mit zwei älteren Angriffsszenarien, die in den Diskursen über Datenschutz und Privatheit der letzten Jahrzehnte jeweils zu ihrer Zeit eine prominente Rolle gespielt haben (siehe zur Übersicht Tab. 1).

*Tab. 1: Qualitativer Vergleich von Angriffsszenarien, die im öffentlichen Diskurs um Datenschutz zu verschiedenen Zeiten eine dominante Bedrohung darstellen. Die jeweils anderen Angriffsszenarien waren zu jeder Zeit ebenfalls denkbar, aufgrund der technologischen Entwicklung besitzt die Relevanz der Szenarien jedoch unterschiedliche zeitliche Schwerpunkte.*

| | Typ 1: Intrusion | Typ 2: Re-Identifikation | Typ 3: Vorhersage |
|---|---|---|---|
| Virulent seit | 1960 ff. | 1990 ff. | 2010 ff. |
| Mittel | Hacking, Datenlecks, Bruch von Verschlüsselung etc. | De-Anonymisierung mittels statistischer Attacken oder Hintergrundwissen | Abschätzung unbekannter Informationen anhand des Abgleichs mit kollektiven Datenbeständen |
| Angriffsziel | persönliche Daten | Anonymität in Datensätzen | Gleichheit der Behandlung, Fairness |
| Schutz | Datensicherheit | Differential Privacy, Federated ML | Predictive Privacy |

### *Typ 1: Intrusion*

Den Urtypus eines Gefahrenszenarios im Datenschutz kann man als Intrusion bezeichnen. Damit eng zusammen hängt die zielgerichtete, auf konkrete Individuen oder Gruppen begrenzte Überwachung. Die Gefahr der gewaltsamen Entwendung von Daten aus mehr oder weniger gesicherten, jedenfalls nicht-öffentlichen Zonen ist tragend für Debatten zum Datenschutz spätestens seit dem Verbreiten der elektronischen Datenverarbeitung in den 1960er Jahren. Das Mittel dieser Form der Verletzungen von Privatheit ist der klassische „Datenklau“ als gezielter Akt der Entwendung

von Daten über technische oder organisatorische Schutzbarrieren hinweg. Obwohl die wichtigste potenzielle Angreiferin immer die datenverarbeitende Organisation selbst ist, wird dieser Angriffstypus in der populären Imagination oft mit *hacking* und Cyberattacken durch Kriminelle oder Geheimdienste in Verbindung gebracht. Das Angriffsziel der intrusiven Verletzung von Privatsphäre sind *konkrete* sensible Datenbestände (über Einzelpersonen, Kohorten, Firmen, staatliche Prozesse, …), die den Angreifenden eigentlich nicht zugänglich sein sollten.

### *Typ 2: Re-Identifikation*

Eine zweiter Angriffstyp wird als Re-Identifikation bezeichnet. Dieser Typus wurde erst in den 1990er Jahren virulent, nachdem durch die Digitalisierung des Gesundheitswesens – zum Beispiel der Abrechnungsvorgänge mit Versicherungen oder der Patientenverwaltung in Krankenhäusern – umfassende digitale Datenbestände über die Prozesse der medizinischen Versorgung verfügbar wurden. Es kam dann die Idee auf, diese Daten für statistische Auswertungen im Rahmen wissenschaftlicher Forschung verwenden zu wollen. Dazu stellte sich die Frage, wie man die Einträge in solchen Datenbanken anonymisieren könnte, um sie dann zu veröffentlichen.

In einem mittlerweile legendären Fall hat der US-Bundesstaat Massachusetts Ende der 1990er Jahre die Krankenhaus-Behandlungsdaten seiner ca. 135.000 staatlichen Bediensteten und ihrer Angehörigen in vermeintlich anonymisierter Form der Forschung zugänglich gemacht. Die Anonymisierung der Datensätze erfolgte, indem Name und Anschrift, sowie die Sozialversicherungsnummer aus den Datensätzen herausgelöscht wurden. Latanya Sweeney, damals Informatik-Studentin am MIT, konnte mit einer linkage-Attacke den Datensatz des damaligen Gouverneurs von Massachusetts, William Weld, in den anonymisierten Daten identifizieren und seine Krankenakte rekonstruieren (Sweeney 2002; Ohm 2010). Dieser Fall hat in Wissenschaft und Politik eine intensive Diskussion über Grenzen und Machbarkeit von Anonymisierung ausgelöst. Die Frage der „sicheren" Anonymisierungsverfahren wird davon ausgehend bis heute diskutiert; jeweils aktuelle Vorschläge für Anonymisierungsverfahren in der Informatik werden immer wieder einige Zeit später durch spektakuläre Angriffe

gebrochen[6]; es ist klar geworden, dass „Anonymität" ein komplexer, nicht absolut definierbarer Begriff ist, der stets von Annahmen in Bezug auf das Hintergrundwissen der Angreifer:in und der statistischen Verteilung der Daten im zu anonymisierenden Datensatz abhängt. Auf Verfahren der Anonymisierung lastet die Anforderung, dass ein heute verwendetes Verfahren alle zukünftigen Angriffstechniken antizipieren und alle möglichen Konfigurationen von Hintergrundwissen zukünftiger Angreifer:innen abdecken muss.[7]

Die Gefahr der Re-Identifizierbarkeit von Individuen in anonymisierten Datensätzen wurde seit den 1990er Jahren zu einem zweiten, viel diskutierten Gefahrenszenario im Datenschutz. Die Diskussion hatte insbesondere spürbaren Einfluss auf die Datenschutzgesetzgebung im Kontext medizinischer Daten, in den USA zum Beispiel auf den *Health Information Portability and Accountability Act* (HIPPA) von 1996. Für die Zwecke des vorliegenden Kapitels kommt es darauf an, auf die qualitative Differenz zum Angriffstyp der Intrusion (und der Prädiktion) hinzuweisen. Im Unterschied zum Datenklau ist das Ziel von Re-Identifikationsattacken ein Bruch der Anonymität. Auch wenn hier ebenfalls sensible Daten über Einzelne oder definierte Kohorten ermittelt werden, ist das etwas anderes als intrusiver Datenklau, da die zugrundeliegenden Daten zuvor bewusst veröffentlicht wurden, jedoch mit dem Versprechen, dabei nichts über Einzelindividuen, sondern nur über statistische Zusammenhänge preiszugeben.

### *Typ 3: Prädiktion*

Mein Argument ist nun, dass auch Re-Identifikation heute schon nicht mehr als der wichtigste und dominante Angriffstypus im Datenschutz gelten kann. Das Prinzip der Vorhersage von unbekannten Daten mittels Big Data und KI-Technologie löst die Gefahr der Re-Identifizierung freilich nicht auf (genauso wenig wie die Gefahr der Intrusion). Die Gefährdung

---

6 Vgl. Ohm 2010 und besonders spektakulär: Die Re-Identifikation von Netflix-Usern in einer pseudonymisiert publizierten Datenbank aus Film-Bewertungen (Narayanan und Shmatikov 2008) oder die Rekonstruktion des Familiennamens anhand anonym vorliegender Genom-Daten (Gymrek u.a. 2013).

7 Die Bundesrepublik Deutschland hat im Dezember 2019 im Rahmen des „Digitale-Versorgung-Gesetz" erst die Zusammenführung der Behandlungsdaten aller ca. 70 Millionen gesetzlich Krankenversicherten zu einer zentralen Forschungsdatenbank beschlossen, vgl. *Bundesgesetzblatt Teil I*, Nr. 49 vom 18.12.2019, S. 2562. Vgl. dazu auch Mühlhoff 2020a.

durch unregulierte prädiktive Analytik übertrifft jedoch beide klassische Angriffsszenarien bei Weitem hinsichtlich Reichweite und Skalierbarkeit. Ist ein prädiktives Modell einmal erstellt – und hierfür gibt es zur Zeit keine wirksamen rechtlichen Beschränkungen –, kann es auf Millionen Nutzer:innen automatisiert und nahezu ohne Grenzkosten angewandt werden. Die Datenfreigiebigkeit der oft privilegierten Gruppe von Nutzer:innen, die vorbehaltlos die Trainingsdaten für prädiktive Analysen bereitstellen (z.B. Gruppe der Facebook-Nutzer:innen, die explizite Angaben über ein sensibles Attribut machen, siehe oben), setzen den Standard des über beinahe *alle* Menschen ermittelbaren Wissens, solange prädiktive Analytik-Technologie nicht reguliert wird.

Dies stellt eine qualitativ neue Gefahrenlage im Datenschutz dar, denn das Mittel der Verletzung prädiktiver Privatheit ist weder der Datenklau noch der Bruch eines Anonymisierungsversprechens. Die Gefährdung durch prädiktive Analytik unterscheidet sich von den älteren Angriffsszenarien in drei Hinsichten: hinsichtlich ihres Ursprungs beruht sie auf der Verfügbarkeit *kollektiver* Datenbestände; hinsichtlich der verübenden Instanz ist sie genau jenen Akteuren vorbehalten, die über aggregierte kollektive Datenbestände verfügen; und hinsichtlich ihrer Effekte zeigt sie nicht allein individuelle sondern vielmehr *gesamtgesellschaftliche* Auswirkungen. Dies bedeutet erstens eine *kommerzielle Zentralisierung* der von prädiktiver Analytik ausgehenden Datenmacht bei wenigen großen Unternehmen. Zweitens liegt der potenzielle Schaden von Verletzungen prädiktiver Privatheit nicht nur in der Abschätzung von Informationen über gezielt ausgewählte Einzelindividuen, sondern in der automatischen und synchronen Abschätzung dieser Informationen über sehr große Nutzer:innen-Kohorten, die eine breite Mehrheit unserer Gesellschaften betreffen. Im Zentrum der Verletzung prädiktiver Privatheit steht also nicht Spionage, die sich auf Einzelne richtet, sondern automatisierte und serienmäßige Ungleichbehandlung von Menschen in der Breite der Gesellschaft. Diese Ungleichbehandlung ist ein *struktureller Faktor* insofern sie sich nicht nur auf Einzelindividuen richtet, sondern auf uns alle in der Interaktion mit automatisierten Systemen, zum Beispiel, wenn uns unterschiedliche Preise für Versicherungen angeboten werden, automatisiert entschieden wird, wer für ein Jobinterview eingeladen wird usf. Das Angriffsziel der Verletzung prädiktiver Privatheit ist somit die Gleichheit und Fairness der gesellschaftlichen Behandlung. Das Schutzgut, das hier verletzt wird, ist im Unterschied zu den anderen Angriffstypen erst in einer kollektivistischen Perspektive erkennbar.